“DIFFERENTIALLY EXPRESSED GENES OF ESCHERECHIA COLI UNDER DIFFERENT STRESSES”

Project report submitted to Sister Nivedita University in partial fulfilment of the requirement for the degree of

MASTER OF SCIENCE

IN

MICROBIOLOGY

Submitted By

MUSHKAN SAHAY

Registration ID: 230010099394

Enrolment ID: 2322211001009

Advisor

Dr. CHITTABRATA MAL

Designation: ASSOCIATE PROFESSOR, Department of Bioinformatics

Maulana Abdul Kalam Azad University of Technology

SNU SISTER NIVEDITA UNIVERSITY

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Utech

In Pursuit Of Knowledge And Excellence

School of Health Science and Translational Research

Sister Nivedita University

June 2025

Table of Contents

1. **Certificate**
2. **Declaration**
3. **Acknowledgement**
4. **List of Figures**
5. **List of Tables**
6. **Abbreviations**
7. **Abstract**
8. **Keywords**
9. Introduction
10. Literature Review
11. Methods and Materials
12. Results
13. Discussion
14. Conclusion
15. Future Prospects
16. References
17. Appendix

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL
Main Campus: NH 12, Haringhata, Post Office - Simhat,
Police Station – Haringhata, Pin - 741249

Date: May 27, 2025

TO WHOM IT MAY CONCERN

This is to certify that Ms. Mushkan Sahay, 2nd year M.Sc. Student (Registration ID: 230010099394, Enrollment ID: 2322211001009) from the Department of Microbiology, Sister Nivedita University, Kolkata, West Bengal, has completed her five months (January - May 2025) dissertation program at the Department of Bioinformatics, Maulana Abul Kalam Azad University of Technology, West Bengal under my guidance on the following topics:

1. *Introduction to NGS Data analysis*
2. *Workflow creation using Python and R programming*
3. *Differential gene expression analysis*
4. *Functional analysis of stress responsive genes*

She has also successfully completed a project titled **"Differentially Expressed Genes of *Escherichia Coli* Under Different Stresses"**.

I found her sincere, hardworking, and enthusiastic, and dedicated in her studies with excellent behavior. Her performance during the internship was highly commendable and I wish her all the best for her future endeavors.

Chittabrata Mal 27/05/2025

Dr. Chittabrata Mal
Associate Professor,
Department of Bioinformatics
Maulana Abul Kalam Azad
University of Technology, West Bengal
Email: chittabrata.mal@makautwb.ac.in

Dr. Chittabrata Mal
Associate Professor
Department of Bioinformatics
Maulana Abul Kalam Azad University
of Technology, West Bengal
(Formerly known as West Bengal University of Technology)
Haringhata, Nadia - 741249, W.B.

SISTER NIVEDITA UNIVERSITY
সিস্টার নিবেদিতা বিশ্ববিদ্যালয়

DG Block - 1/2, Action Area I, New Town, Kolkata - 700156, West Bengal, India.

DEPARTMENT OF BIOTECHNOLOGY & DEPARTMENT OF MICROBIOLOGY
জৈব প্রযুক্তি বিভাগ ও অণুজীব বিদ্যা বিভাগ

TO WHOM IT MAY CONCERN

This is to certify that the project report titled "**DIFFERENTIALLY EXPRESSED GENES OF *ESCHERICHIA COLI* UNDER DIFFERENT STRESSES** ", submitted by **Ms. Mushkan Sahay (Registration No. 230010099394, Enrolment No. 2322211001009)** to **the Department of Microbiology**, School of Health Science and Translational Research at Sister Nivedita University, Kolkata, is a bona fide record of research conducted by him/her for the partial fulfillment of the requirements for the Master of Science degree.

Date: 05.06.2025

Place: Newtown, Kolkata

Signature of Head of the Department

DECLARATION BY THE STUDENT

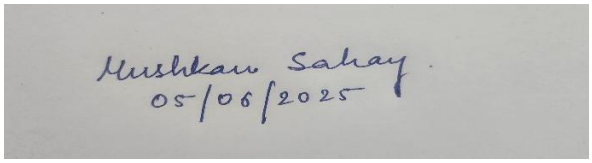
I hereby declare that the work which is being presented in this dissertation entitled "**Differentially Expressed Genes of Escherichia Coli Under Different Stresses**"

in partial fulfilment of the requirements for the award of the degree of Masters of Science in Microbiology in is an authentic record of my own work carried out during the period from 15th January to 30th May under the supervision of Dr. Chittabrata Mal, Associate Professor, Department of Bioinformatics, Maulana Abdul Kalam Azad University of Technology, West Bengal

I acknowledge that I have read and understood the UGC regulations, 2018 (*Promotion of Academic integrity and Prevention of Plagiarism in Higher Educational Institutions*).

I confirm that this dissertation has been checked for plagiarism using the plagiarism checking software "SmallSEOTools". At the end of the dissertation, a copy of the summary report demonstrating similarities in content and its potential source (if any), generated online using plagiarism checking software is enclosed. I herewith confirm that the dissertation has been less than 10% similarity according to the plagiarism checking software's report and meets the UGC Regulations as well as plagiarism rules of Sister Nivedita University.

I further declare that no portion of the dissertation or its data will be published without the Institute's or supervisor's permission. I have not previously applied for any other degree or award using the topics and findings described in my dissertations.



(Signature of the students with Date)

Name of the Student: Mushkan Sahay

Registration no: 230010099394

Department: Microbiology

Acknowledgement

I would like to express my deepest gratitude to my supervisor, **Dr. Chittabrata Mal**, for his unwavering support, insightful guidance, and continuous encouragement throughout this research journey.

As a student transitioning from a microbiology background into the interdisciplinary realm of transcriptomic data analysis, I faced several learning challenges. Dr. Chittabrata Mal was exceptionally understanding of these difficulties and consistently provided me the time and space to grasp complex concepts at my own pace. His patience, empathy, and flexibility made a significant difference in helping me adapt and grow in this new field.

Dr. Chittabrata Mal's constructive feedback, academic expertise, and dedication to excellence have been instrumental in shaping the quality and direction of this thesis. More than just a mentor, their encouragement helped me stay motivated and confident through the ups and downs of research.

I would also like to sincerely thank **Sister Nivedita University** and all the faculty members and professors of the School of Health Sciences and Translational Research for their consistent support, encouragement, and academic guidance throughout my postgraduate journey. Their efforts have provided a strong foundation for my academic and professional development.

It has been an honour and a privilege to work under Dr. Chittabrata Mal's mentorship and be a part of Sister Nivedita University. I will always remain grateful for the knowledge, values, and research ethics I have gained during this journey.

List of Figures

S. No.	Figure Number	Description
1	Figure 4.1.1 (a)	Bar graph showing average log2 fold change of upregulated genes in GSE181239
2	Figure 4.1.1 (b)	Boxplot showing distribution of log2 fold change values for upregulated genes (GSE181239)
3	Figure 4.1.2 (a)	Bar graph showing average log2 fold change of downregulated genes in GSE181239
4	Figure 4.1.2 (b)	Boxplot showing distribution of log2 fold change values for downregulated genes (GSE181239)
5	Figure 4.2.1 (a)	Bar graph showing average log2 fold change of upregulated genes in GSE221524
6	Figure 4.2.1 (b)	Boxplot showing distribution of log2 fold change for upregulated genes (GSE221524)
7	Figure 4.2.2 (a)	Bar graph showing average log2 fold change of downregulated genes in GSE221524
8	Figure 4.2.2 (b)	Boxplot showing distribution of log2 fold change of downregulated genes in GSE221524
9	Figure 4.3.1 (a)	Bar graph showing average log2 fold change of upregulated genes in GSE208658
10	Figure 4.3.1 (b)	Boxplot showing distribution of log2 fold change of upregulated genes in GSE208658
11	Figure 4.3.2 (a)	Bar graph showing average log2 fold change of downregulated genes in GSE208658
12	Figure 4.3.2 (b)	Boxplot showing distribution of log2 fold change of

		downregulated genes in GSE208658
13	Figure 4.4.3	Bar chart showing genes with highest frequency of upregulation across samples
14	Figure 5.2.1	Log2 fold change of mqsR across stress conditions
15	Figure 5.2.2	Log2 fold change of bhsA across 13 stress conditions
16	Figure 5.2.3	Log2 fold change of ygiS across 13 stress conditions
17	Figure 5.2.4	Log2 fold change of fadE across 13 conditions
18	Figure 5.2.5	Log2 fold change of pspA across 13 downregulated conditions
19	Figure 5.2.6	Log2 fold change of nhaA across all 13 samples
20	Figure 5.2.7	Log2 fold change of nanT across all 13 samples
21	Figure 5.2.8 (a)	Bar plot comparing log2 fold change of seven downregulated genes (mqsR, bhsA, ygiS, fadE, pspA, nhaA, nanT) across datasets GSE208658, GSE181239, and GSE221524
22	Figure 5.2.8 (b)	Heatmap showing expression patterns of downregulated genes across stress types
23	Figure 5.2.8 (c)	Functional grouping of downregulated genes

List of Tables

Table No.	Title
Table 4.1	Summary of Upregulated Genes Across Nickel and Genetic Stress Comparisons in E. coli
Table 4.1.2	Summary of Downregulated Genes Across Nickel and Genetic Stress Comparisons in E. coli
Table 4.2.1	Upregulated Gene Expression Summary for Protein Secretion Stress Conditions (GSE221524)
Table 4.2.2	Downregulated Gene Expression Summary for Protein Secretion Stress Conditions (GSE221524)
Table 4.3.1	Upregulated Gene Expression Summary Across Radiation Stress Comparisons (GSE208658)
Table 4.3.2	Downregulated Gene Expression Summary Across Radiation Stress Comparisons (GSE208658)
Table 4.4.1	Genes Ranked Based on Appearance Frequencies Across Samples (Upregulated)
Table 4.4.2	Frequency-Based Gene Filtering for Upregulated Genes
Table 4.5.1	Genes Ranked Based on Appearance Frequencies Across Samples (Downregulated)
Table 4.5.2	Bar Chart Summary of Top 8 Most Frequently Downregulated Genes (Used as Figure, but contains table-like data)
Table 4.6	Summary of All the Upregulated and Downregulated Genes
Table 4.7.1	Phylogenetic Grouping of Upregulated Genes
Table 4.7.2	Phylogenetic Grouping of Downregulated Genes
Table 5.1.10	Functional Annotation of Upregulated Genes
Table 5.2.9	Functional Annotations and Pathway Mapping of Downregulated Genes in E. coli Under Stress Conditions

List of Abbreviations

Abbreviations	Full form
ALE	Adaptive Laboratory Evolution
ATP	Adenosine Triphosphate
DAVID	Database for Annotation, Visualization and Integrated Discovery
DEGs	Differentially Expressed Genes
DESeq2	Differential Expression Sequencing 2 (R package)
DGE	Differential Gene Expression
DNA	Deoxyribonucleic Acid
FC	Fold Change
FDR	False Discovery Rate
GEO	Gene Expression Omnibus
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria-Bertani (medium)
log2FC	Log2 Fold Change
MSA	Multiple Sequence Alignment
mRNA	Messenger RNA
NCBI	National Center for Biotechnology Information
$NiCl_2$	Nickel (II) Chloride
PCA	Principal Component Analysis
PBS	Phosphate Buffered Saline (implied in wet-lab contexts)
RNA	Ribonucleic Acid
RNA-Seq	RNA Sequencing
ROS	Reactive Oxygen Species
Rpm	Reads Per Million (in expression normalization)
tRNA	Transfer RNA
TGT	tRNA-Guanine Transglucosylase
WT	Wild Type
Δtgt	tgt Gene Knockout Mutant

Abstract

Understanding how *Escherichia coli* (E. coli) reacts to different kinds of stress at the molecular level is important for both basic microbiology and applications in biotechnology. In this study, I looked at how E. coli's gene expression changes under three types of stress: oxidative stress caused by radiation, heavy metal stress from nickel exposure, and genetic stress due to deletion of the *tgt* gene, which affects tRNA modification. For this analysis, I used three publicly available RNA-Seq datasets (GSE181239, GSE221524, and GSE208658) and performed differential gene expression analysis using the DESeq2 software (Love et al., 2014). In total, 22 pairwise comparisons were carried out to find significantly regulated genes under each condition.

From the analysis, I found eight genes (*zntA*, *rlmN*, *lysS*, *gltX*, *accC*, *gltB*, *serC*, and *yebC*) that were repeatedly upregulated under different stress situations. These genes are known to be involved in important functions like metal detoxification, protein translation, and amino acid biosynthesis (Macomber & Hausinger, 2011; Fergus et al., 2015; Zhao & Winkler, 1996; Chugani et al., 2001). On the other hand, seven genes (*mqsR*, *bhsA*, *ygiS*, *fadE*, *pspA*, *nhaA*, and *nanT*) were commonly downregulated and are generally linked to stress signalling, maintaining cell membrane integrity, or energy-intensive processes (Yamaguchi et al., 2011; Darwin, 2005). This suggests that E. coli might shut down some processes to save energy during long periods of stress.

To better understand which responses were general versus specific to a stress type, I used a frequency-based comparison method. Genes seen in multiple datasets were marked as core stress response genes, while others were considered condition-specific. I also used heatmaps and clustering to visually show patterns of gene expression and co-regulation.

Overall, this project provides insights into how E. coli adapts to various stressors at the transcriptome level. The genes identified here could be useful in future research focused on stress tolerance or bioengineering. This work adds to our knowledge of bacterial survival strategies and may help in developing more resilient microbial strains for industrial or scientific use.

Keywords

Escherichia coli, transcriptomics, RNA-Seq, stress response, differentially expressed genes, oxidative stress, nickel toxicity, *tgt* deletion, gene regulation, DESeq2, adaptive mechanisms, microbial biotechnology.

spe

1. Introduction

1.1. E. coli as a Model Organism

Escherichia coli (*E. coli*) is a Gram-negative, rod-shaped bacterium commonly present in the digestive systems of humans and other mammals. Due to its quick reproduction, fully sequenced genome, and relative simplicity in genetic manipulation, *E. coli* has become a central model organism widely employed in microbiology, molecular biology, and biotechnological research (Blattner et al., 1997). Researchers regularly use *E. coli* to study fundamental biological mechanisms including DNA synthesis, gene expression control, and metabolic responses.

1.2. Importance of Studying Stress Responses through Transcriptomics

In both laboratory cultures and natural settings, bacteria such as *E. coli* frequently experience different types of stress, such as oxidative damage, exposure to heavy metals, disruptions in membrane integrity, and internal genetic alterations. To cope with these challenges, bacterial cells initiate carefully orchestrated changes in gene expression, enabling adaptation and survival. Analyzing these stress-induced transcriptional changes using transcriptomics provides a comprehensive understanding of the global regulatory networks and metabolic pathways activated under stressful conditions (Wang et al., 2009).

RNA sequencing (RNA-Seq) has emerged as a powerful method to measure gene expression at a genome-wide scale with remarkable sensitivity and accuracy. Compared to traditional techniques like microarrays, RNA-Seq can detect a broader dynamic range of gene expression, including previously unknown or rare transcripts, making it especially effective for studying complex bacterial stress responses.

1.3. RNA-Seq and Differential Gene Expression (DGE) Analysis

RNA-Seq methodology begins with extracting RNA from biological samples, converting it to complementary DNA (cDNA), fragmenting it, and sequencing these fragments using advanced sequencing technologies. The resulting short reads are mapped against a known reference genome, enabling quantification of expression levels for each gene. Differential Gene Expression (DGE) analysis then identifies significant expression changes between different experimental conditions or treatment groups.

DESeq2 is a popular computational tool used for DGE analysis, employing statistical models based on the negative binomial distribution to effectively handle biological variability and sequencing depth differences (Love et al., 2014). DESeq2 outputs include values like log2 fold changes and adjusted p-values, highlighting the specific genes significantly regulated under various conditions.

1.4 Rationale for Investigating Specific Stress Conditions

The current research examines how *E. coli* modifies its transcriptome under four specific stress scenarios, using existing RNA-Seq datasets to represent biologically relevant conditions:

- **Metal stress:** Studied using *E. coli* strain K-12 MG1655 exposed to 2 mM nickel chloride ($NiCl_2$), this condition models the cellular response mechanisms triggered by toxic heavy metals (GEO accession GSE181239).
- **Genetic stress:** Examined via deletion of the *tgt* gene, responsible for queuosine modification in tRNAs, highlighting the transcriptomic consequences of disrupted translation accuracy and efficiency (GEO accession GSE181239).
- **Oxidative and genotoxic stress:** Induced by acute and chronic radiation exposure from isotopes (Pu-239, Fe-55, and H-3), modeling how cells manage DNA and macromolecular damage resulting from oxidative challenges (GEO accession GSE208658).
- **Envelope stress:** Investigated using *E. coli* strain Nissle 1917 engineered to secrete proteins via the curli system, reflecting the cellular stress associated with protein overexpression, secretion, and membrane transport burdens (GEO accession GSE221524).

These conditions were deliberately chosen because they represent diverse yet interconnected physiological stresses, enabling a comprehensive exploration of both shared and condition-specific gene regulatory pathways in bacterial stress responses.

1.5. Study Objectives

This thesis aims to investigate and compare the transcriptomic adaptations of *E. coli* subjected to multiple forms of environmental and genetic stress. By analyzing three distinct RNA-Seq datasets with a uniform DESeq2 workflow, the study specifically intends to:

- Identify significantly regulated genes (both upregulated and downregulated) in response to metal toxicity, genetic deletion, radiation-induced oxidative stress, and protein secretion stress.
- Examine overlaps and distinctions among gene expression patterns across these varied stress conditions.
- Conduct functional enrichment analyses to identify the biological pathways critical for bacterial adaptation and survival.
- Highlight key genes or regulatory networks that consistently participate in stress responses, potentially serving as biomarkers for future research.
- Establish a robust and integrated transcriptomic reference that supports subsequent experimental validation or applied biotechnology projects.

2. Literature Review

2.1 Transcriptomic Research in *E. coli*

Transcriptomic techniques have become key tools for understanding how *Escherichia coli* responds to different environmental and physiological stressors. Initially, researchers relied on microarray technology to study global gene expression, but these approaches had limitations in sensitivity and dynamic range. The introduction of RNA sequencing (RNA-Seq) provided a more powerful alternative, capable of detecting low-abundance transcripts and novel regulatory RNAs (Conway et al., 2014).

Recent studies have widely applied RNA-Seq to investigate how *E. coli* responds to stressors such as nutrient depletion, heat or cold shock, shifts in pH, and antibiotic pressure. One common observation is the involvement of transcriptional regulators like OxyR and SoxRS, which initiate stress response pathways by activating genes that protect the cell from damage caused by reactive oxygen species (ROS) (Greenberg & Demple, 1989; Seo et al., 2015).

2.2 Queuosine Function and the *tgt* Gene

Queuosine is a unique modification added to certain tRNAs at the wobble position, affecting amino acids like histidine, aspartic acid, asparagine, and tyrosine. This modification is introduced by the enzyme tRNA-guanine transglycosylase, which is encoded by the *tgt* gene (Noguchi & Nishimura, 1982). When the *tgt* gene is deleted, the affected tRNAs lack queuosine, leading to reduced translation fidelity and heightened sensitivity to stress. Transcriptomic studies show that *tgt* mutants exhibit notable gene expression changes, especially under stressors like metal toxicity and oxidative damage, suggesting a broader protective role for queuosine in cellular homeostasis (Fergus et al., 2015).

2.3 Oxidative and Radiation-Induced DNA Stress

Oxidative stress arises when the balance between ROS production and the cell's detoxification capacity is disturbed. ROS like hydrogen peroxide and superoxide can damage vital components such as DNA, proteins, and membranes. In *E. coli*, transcription factors including OxyR and SoxRS activate defense mechanisms that involve antioxidant and DNA repair systems (Imlay, 2013).

In this project, one of the key datasets used (GSE208658) involves *E. coli* exposed to radioactive elements including Pu-239, Fe-55, and H-3. These isotopes induce both short- and long-term oxidative damage. Comparing 1-day versus 15-day exposure durations reveals how the bacterium modifies its transcriptome to adapt to persistent oxidative stress, providing insights into long-term stress management strategies (Rousset et al., 2021).

2.4 Nickel Stress and Cellular Responses

Nickel, while essential in trace amounts for certain enzymes, becomes harmful when present in excess. High levels interfere with enzymatic functions and promote oxidative damage. *E. coli* has developed detoxification systems to counteract nickel toxicity. These include efflux pumps like *rcnA* and regulatory proteins such as *rcnR* that help control intracellular metal concentrations (Rodrigue et al., 2005).

RNA-Seq data from the nickel stress experiment (GSE181239) shows considerable changes in genes related to metal ion homeostasis, oxidative defense, and membrane stabilization. These responses reflect how the bacterium adapts to metal-induced challenges and could be useful in industrial applications like bioremediation.

2.5 Stress Linked to Protein Export and Envelope Disruption

Overproduction and secretion of proteins can place stress on the bacterial envelope, especially when the system is overloaded. One model for studying this is the *E. coli* Nissle 1917 strain, engineered to export proteins via the curli pathway. The dataset (GSE221524) used in this thesis focuses on this kind of envelope stress.

When secretion demand increases, cells activate genes involved in membrane maintenance, protein folding (via chaperones), and transport systems to prevent damage. Studying such responses not only enhances our understanding of bacterial adaptation but also supports the design of more efficient strains for industrial protein production.

3. Methods and Materials

3.1 Data Acquisition

To explore how *Escherichia coli* (*E. coli*) modifies its transcriptome under various stress conditions, three RNA sequencing (RNA-Seq) datasets were obtained from the NCBI Gene Expression Omnibus (GEO) database. These datasets—GSE181239, GSE221524, and GSE208658—represent distinct biological stress scenarios, including genetic alterations, environmental pressures, and radiation exposure, thus enabling robust comparative transcriptomic analysis.

- **Dataset GSE181239** involves experiments on *E. coli* K-12 MG1655 to investigate how nickel stress and the deletion of the *tgt* gene (which impairs queuosine modification of tRNAs) influence gene expression. In this experiment, both wild-type and Δ*tgt* mutant strains were cultured in LB medium, with or without the addition of 2 mM nickel chloride.
- **Dataset GSE221524** examines the transcriptomic impact of protein expression and secretion stress in engineered *E. coli* Nissle 1917 strains. These strains were designed either to produce proteins internally or export them via the curli secretion pathway, triggered by IPTG induction.
- **Dataset GSE208658** assesses how chronic exposure to low-dose radiation (isotopes Pu-239, Fe-55, and H-3) affects the transcriptome of *E. coli* DH10β. Experiments included both short-term (1-day) and long-term (15-day) exposures, alongside chemically matched, non-radioactive control samples.

These datasets were specifically chosen to ensure a wide representation of stress types, allowing systematic comparison of gene regulatory responses across diverse stress conditions.

3.2 Sample Metadata and Experimental Design

Dataset GSE181239 utilized a factorial design (2×2) involving two genetic backgrounds (wild-type and Δ*tgt*) and two environmental conditions (with or without nickel). Each condition was replicated three times, resulting in 12 total samples. The experimental design enabled the following key transcriptomic comparisons:

- Wild-type with nickel vs. wild-type control: identifying nickel-induced gene changes.
- Δ*tgt* mutant vs. wild-type under normal conditions: assessing the impact of losing queuosine modification at baseline.
- Δ*tgt* mutant with nickel vs. wild-type with nickel: understanding how the absence of queuosine affects nickel-induced gene expression.
- Δ*tgt* mutant with nickel vs. Δ*tgt* without nickel: evaluating the specific response of the mutant strain to nickel stress.

Dataset GSE221524 involved three genetically engineered *E. coli* Nissle 1917 strains designed to express recombinant proteins under controlled conditions:

- "All-Tags" without IPTG (0 μM): baseline control with no induced protein production.

- "All-Tags" with IPTG (10 μM): protein expression with secretion through the curli system.
- "No-Sec" with IPTG (10 μM): protein expressed internally without secretion.

Each of these conditions included three biological replicates, totaling nine samples. This design permitted direct comparison to isolate the effects associated with protein production and secretion stress on bacterial gene expression.

Dataset GSE208658 explored *E. coli* DH10β's response to ionizing radiation from isotopes Pu-239, Fe-55, and H-3. Experiments were conducted over short-term (1 day) and extended (15 days) radiation exposures, each compared against chemically identical but non-radioactive controls. Each condition contained three replicates, allowing robust statistical comparisons. In total, 15 distinct pairwise comparisons were defined:

1. Pu239 exposure (1 day) vs. control (1 day)
2. Pu239 exposure (15 days) vs. control (15 days)
3. Pu239 exposure (15 days) vs. Pu239 exposure (1 day)
4. Fe55 exposure (1 day) vs. control (1 day)
5. Fe55 exposure (15 days) vs. control (15 days)
6. Fe55 exposure (15 days) vs. Fe55 exposure (1 day)
7. H3 exposure (1 day) vs. control (1 day)
8. H3 exposure (15 days) vs. control (15 days)
9. H3 exposure (15 days) vs. H3 exposure (1 day)
10. Pu239 exposure (1 day) vs. H3 exposure (1 day)
11. Pu239 exposure (1 day) vs. Fe55 exposure (1 day)
12. Fe55 exposure (1 day) vs. H3 exposure (1 day)
13. Pu239 exposure (15 days) vs. H3 exposure (15 days)
14. Pu239 exposure (15 days) vs. Fe55 exposure (15 days)
15. H3 exposure (15 days) vs. Fe55 exposure (15 days)

These comparisons facilitated a comprehensive investigation of transcriptional changes, distinguishing both the effects of exposure duration and the distinct impacts of specific radioactive isotopes. Consequently, this dataset provides valuable insights into bacterial adaptation mechanisms during chronic radiation stress.

3.3 Data Preprocessing

Raw count matrices for each dataset were downloaded from the NCBI Gene Expression Omnibus (GEO) and preprocessed using R version 4.3.1 along with relevant Bioconductor packages. All preprocessing steps were designed to ensure data quality, consistency, and compatibility with the DESeq2 differential expression analysis framework.

- **Data Import:** Raw count data were loaded using the `read.csv()` function and converted into integer matrices with appropriate sample identifiers set as row and column names. Sample column names were programmatically matched with their corresponding experimental metadata to ensure correct pairing.
- **Gene Filtering:** To eliminate noise from low-abundance transcripts, genes were excluded if their total read count was below 10 across all samples or if they were

expressed in fewer than 3 samples. This step ensured that downstream analysis was focused on genes with sufficient expression signals.

- **Normalization:** Library size differences and sequencing depth variability were addressed using DESeq2's default median-of-ratios normalization method. This approach is robust to extreme values and accounts for compositional differences across samples.
- **Transformation for Visualization:** For exploratory data analysis and dimensionality reduction, a variance-stabilizing transformation (VST) was applied to the normalized count data. This transformation reduces the mean-dependent variance typically observed in count data, making it suitable for PCA and clustering.
- **Quality Control:** Principal component analysis (PCA) and hierarchical clustering were performed to visually assess data structure and identify any potential outliers or batch effects. All samples clustered according to biological condition as expected, and no outliers were detected.

3.4 Differential Gene Expression (DGE) Analysis

Differential gene expression analysis was conducted using the DESeq2 pipeline (Love et al., 2014), which models count data using a negative binomial distribution and performs shrinkage-based estimation of dispersion and fold change.

The following steps were carried out for each dataset:

- **DESeq Dataset Creation:** Count matrices and matched sample metadata were used to construct a DESeqDataSet object using the DESeqDataSetFromMatrix() function. The design formula ~ condition was specified to model the effect of experimental conditions on gene expression.
- **Model Fitting and Dispersion Estimation:** The DESeq() function was applied to perform size factor estimation, estimate dispersion values for each gene, and fit the negative binomial model.
- **Contrast Definition and Extraction:** Differential expression results were extracted using the results() function with user-defined contrast vectors (e.g., c("condition", "Treatment", "Control")) to compare specific experimental groups.
- **Significance Filtering:** Genes were considered differentially expressed if they met both of the following criteria: (1) an adjusted p-value (Benjamini-Hochberg FDR) less than 0.05 and (2) an absolute $\log_2$ fold change greater than 1.
- **Output and Storage:** All filtered results were exported as .csv files for each comparison and used in subsequent visualization and interpretation workflows, including heatmaps, volcano plots, and pathway enrichment analyses.

3.5 Gene Annotation

Gene annotation was performed to map locus-specific identifiers to meaningful gene symbols, enabling downstream biological interpretation.

- For **GSE221524**, which involved *E. coli* Nissle 1917, gene identifiers followed the format "gene-CIW80_RSXXXXX." A genome-specific GFF3 annotation file was used, and a custom R script parsed ID= and Name= attributes to associate each feature with its corresponding gene name (e.g., gene-CIW80_RS00010 → hrpA).

- For **GSE181239** and **GSE208658**, most gene symbols were already present in the count matrices. When required, additional mapping to standardized gene names or Entrez IDs was performed using Bioconductor's org.EcK12.eg.db package and the bitr() function from clusterProfiler.

This uniform annotation strategy ensured compatibility across datasets and consistency in downstream functional enrichment and comparative analyses.

3.6 Multiple Sequence Alignment (MSA) and Phylogenetic Analysis

To investigate sequence-level relationships among the most significantly upregulated and downregulated genes identified across datasets, a **multiple sequence alignment (MSA)** was performed using **Clustal Omega**, a widely used alignment tool for protein sequences.

- **Protein sequence retrieval**: Amino acid sequences corresponding to the differentially expressed genes were retrieved from NCBI or UniProt in FASTA format.
- **Alignment**: Sequences were aligned using Clustal Omega with default parameters, generating multiple sequence alignment files.
- **Interpretation**: The resulting alignments were analyzed for conserved residues and patterns of similarity. For the set of downregulated genes, the alignment spanned several hundred residues with **0% fully conserved positions**, indicating a high degree of sequence variability and a lack of shared evolutionary origin. Similar observations were made for the upregulated gene set, suggesting that the genes represent **functionally diverse proteins** involved in **non-redundant stress response mechanisms**.

To further assess the evolutionary relationships among these genes:

- A **phylogenetic tree** was generated from the aligned sequences using Clustal Omega's default tree-building algorithm (neighbor-joining).
- The phylogenetic analysis did not reveal strong clustering among genes based on stress condition or functional category, implying **divergent evolutionary trajectories**

4. **RESULTS**

4.1. GSE181239

4.1.1. Upregulated genes

Comparison	No. of Upregulated Genes	Average log2 Fold Change	Average Adjusted P-value
WT_Ni vs WT_LB	824	1.897	0.00654
WT_LB vs tgt_LB	181	1.843	0.00954
tgt_Ni vs WT_Ni	523	1.136	0.0102
tgt_Ni vs tgt_LB	691	1.532	0.00743

Table : 4.1. *Summary of Upregulated Genes Across Nickel and Genetic Stress Comparisons in E. coli*

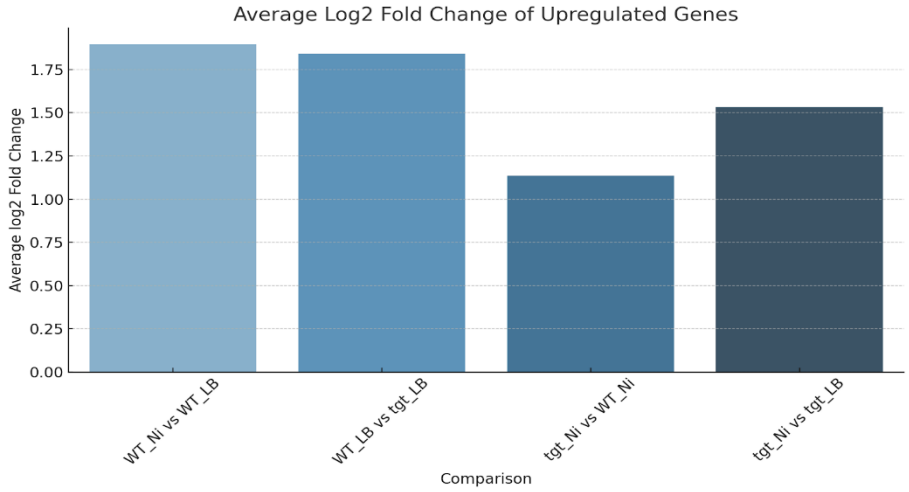


Fig : 4.1.1 (a) : Bar graph showing average log2 fold change of upregulated genes in GSE181239

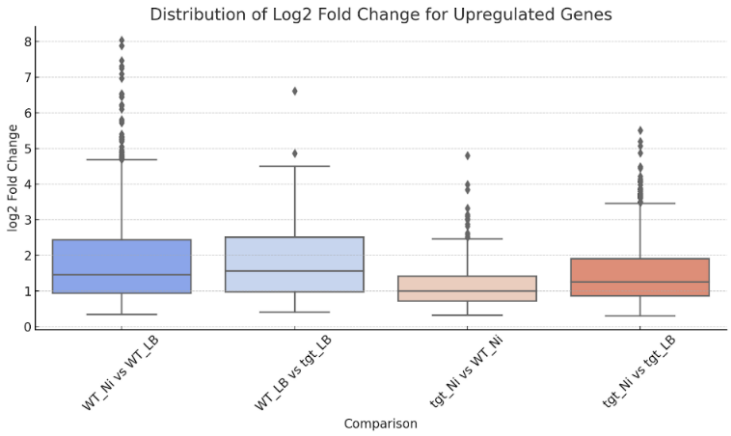


Fig : 4.1.1 (b) Boxplot showing distribution of log2 fold change values for upregulated genes (GSE181239)

Interpretation of Upregulated Genes

The transcriptomic analysis of dataset GSE181239 revealed strong upregulation of genes under both nickel exposure and *tgt* gene deletion. In the comparison between wild-type cells grown in nickel-supplemented medium (WT_Ni) and those in standard LB medium (WT_LB), 824 genes were significantly upregulated, with an average $\log_2$ fold change of 1.897 and an adjusted p-value below 0.007. This robust gene activation suggests a major shift in transcriptional programs when *E. coli* encounters nickel-induced stress.

Similarly, deletion of the *tgt* gene (WT_LB vs tgt_LB) led to upregulation of 181 genes, indicating that even in the absence of an external stressor, loss of queuosine modification can trigger a significant transcriptional response. The *tgt* gene is known to encode tRNA-guanine transglycosylase, which modifies specific tRNAs to maintain translational fidelity (Fergus et al., 2015; Durand et al., 1997).

In comparisons involving the mutant strain under nickel stress (tgt_Ni vs WT_Ni and tgt_Ni vs tgt_LB), hundreds of genes also showed increased expression, indicating that both genetic and environmental stressors independently drive substantial gene regulation.

Functional investigation of the upregulated genes revealed the involvement of key stress-adaptive pathways. The *zntA* gene, a well-characterized metal efflux pump, was among the most significantly induced. Its product helps reduce intracellular nickel levels by actively transporting metal ions out of the cell, thus protecting against toxicity (Macomber & Hausinger, 2011). Also upregulated were genes such as *rlmN*, involved in ribosomal RNA methylation, and aminoacyl-tRNA synthetases like *lysS* and *gltX*, which play critical roles in ensuring translational accuracy (Fergus et al., 2015).

These results collectively highlight how *E. coli* mobilizes distinct transcriptional responses to cope with metal toxicity and internal genetic perturbations. The upregulated genes not only reflect immediate protective measures but may also support long-term adaptation by stabilizing essential cellular processes like protein synthesis and detoxification.

4.1.2 Downregulated

Comparison	No. of Downregulated Genes	Average log2 Fold Change	Average Adjusted P-value
WT_Ni vs WT_LB	662	-1.864	0.00637
WT_LB vs tgt_LB	393	-1.003	0.01199
tgt_Ni vs WT_Ni	503	-1.406	1.00713
tgt_Ni vs tgt_LB	658	-1.411	0.00707

Table 4.1.2 – *Summary of Downregulated Genes Across Nickel and Genetic Stress Comparisons in E. coli*

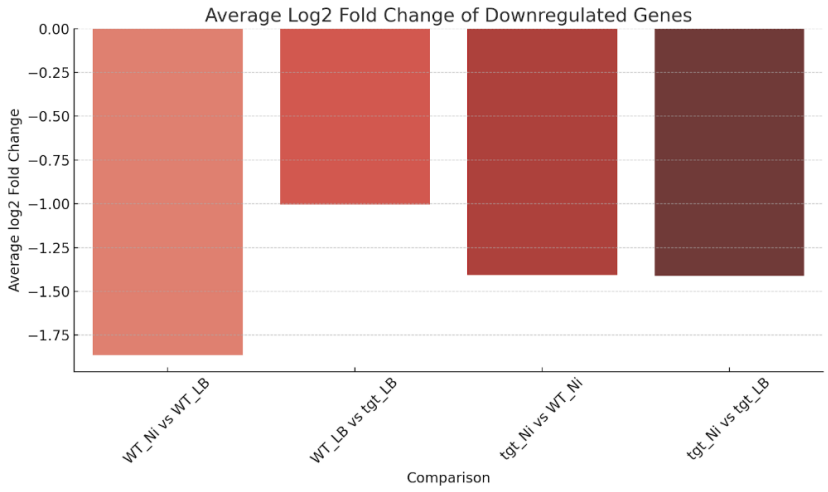


Fig : 4.1.2 (a) Bar graph showing average log2 fold change of downregulated genes in GSE181239

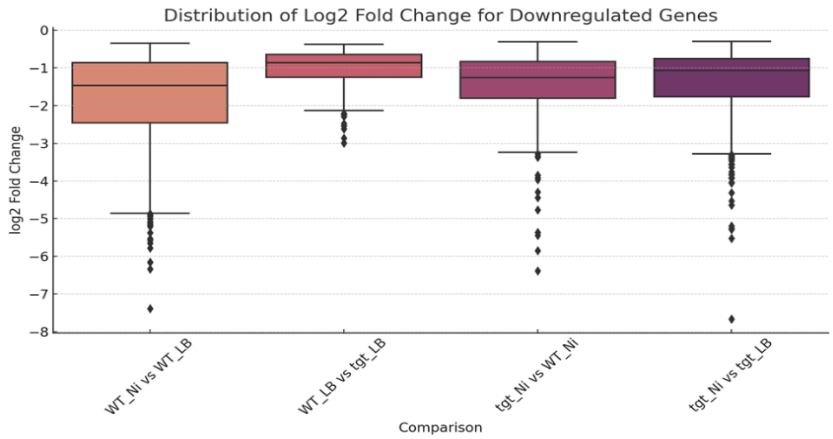


Fig : 4.1.2 (b) Boxplot showing distribution of log2 fold change values for downregulated genes (GSE181239)

Interpretation of Downregulated Genes

In terms of transcriptional repression, the **WT_Ni vs WT_LB** comparison also exhibited the strongest response, with **662 genes** downregulated and an average log2 fold change of **-1.864**. This reflects a broad reduction in gene activity as a response to nickel toxicity. The comparisons **tgt_Ni vs WT_Ni** and **tgt_Ni vs tgt_LB** revealed similarly high levels of downregulation (503 and 658 genes, respectively), with average log2 fold changes around **-1.4**, suggesting that the *tgt* mutant experiences a common stress-repressive transcriptional profile under nickel conditions. In contrast, the **WT_LB vs tgt_LB** comparison showed fewer downregulated genes (**393**) and a milder fold change (**-1.003**), indicating a more targeted response due to *tgt* deletion alone. These findings are consistent with previous reports that stress adaptation in *E. coli* often involves both the upregulation of defense

mechanisms and the downregulation of non-essential or energetically expensive processes (Imlay, 2013). This dual regulation allows the cell to conserve resources while maximizing survival under unfavorable conditions.

4.2 GSE221252

4.2.1. UPREGULATED

Comparison	No. of Upregulated Genes	Average log2 Fold Change	Average Adjusted P-value
All-Tags_10uM vs All-Tags_0uM	670	1.173	0.01001
No-Sec_10uM vs All-Tags_0uM	12	2.115	0.00311
No-Sec_10uM vs All-Tags_10uM	863	1.187	0.00927

Table 4.2.1 – *Upregulated Gene Expression Summary for Protein Secretion Stress Conditions (GSE221252)*

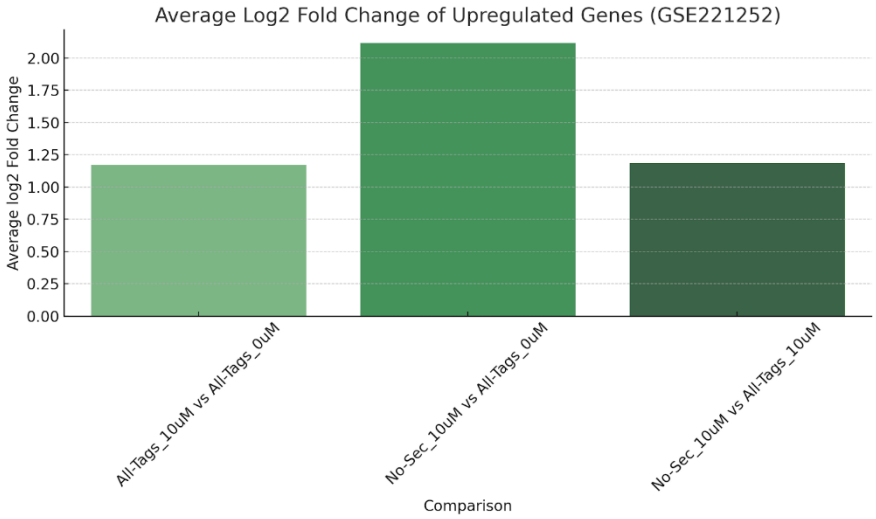


Fig : 4.2.1 (a) Bar graph showing average log2 fold change of upregulated genes in GSE221524

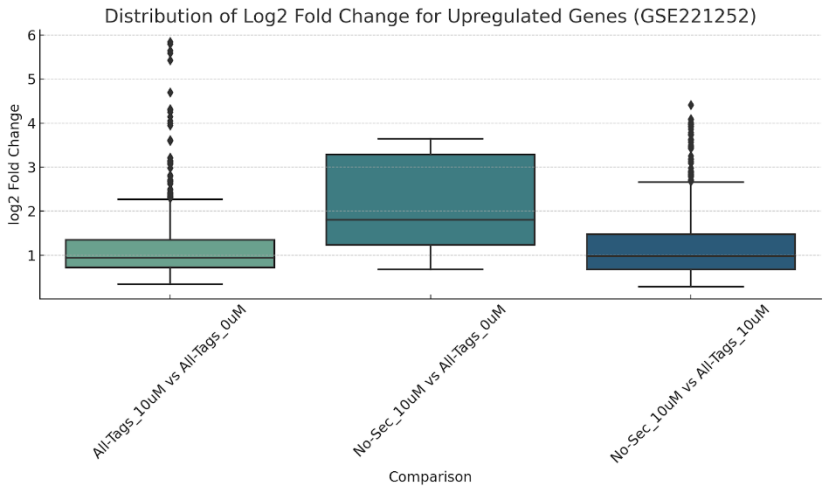


Fig : 4.2.1 (b) Boxplot showing distribution of log2 fold change for upregulated genes

INTERPRETATION

The transcriptomic analysis of *Escherichia coli* Nissle 1917 under protein overexpression and secretion stress revealed distinct patterns of gene upregulation across three experimental comparisons. In the **All-Tags_10uM vs All-Tags_0uM** comparison, a total of **670 genes** were significantly upregulated, with an average log2 fold change of **1.173**. This suggests that simultaneous protein expression and secretion via the curli system triggers a broad transcriptional response, likely due to the energetic and membrane-associated burden of secretion stress (Borkowski et al., 2016).

In contrast, the **No-Sec_10uM vs All-Tags_0uM** comparison showed only **12 upregulated genes**, albeit with a higher average log2 fold change of **2.115**. These results indicate that expression alone, without secretion, elicits a more targeted but stronger response per gene, possibly involving stress sensors that detect cytoplasmic accumulation of proteins (Costa et al., 2014).

The **No-Sec_10uM vs All-Tags_10uM** comparison had the **highest number of upregulated genes (863)**, reflecting significant differences between secretion-competent and secretion-deficient strains. The moderate average log2FC (**1.187**) suggests widespread upregulation of genes involved in envelope maintenance, protein folding, and transport machinery, which are commonly activated in response to secretion stress (Raivio, 2005).

Together, these findings demonstrate that both secretion and expression stresses independently alter gene expression in *E. coli*, with the greatest transcriptomic shift observed when comparing secretory vs non-secretory strains. This supports prior evidence that the bacterial envelope stress response is tightly linked to the protein export burden, and that differential regulation enables the cell to manage folding, trafficking, and integrity of exported proteins (Silhavy et al., 2010).

4.2.2 DOWNREGULATED

Comparison	No. of Upregulated Genes	Average log2 Fold Change	Average Adjusted P-value
All-Tags_10uM vs All-Tags_0uM	678	-1.140	0.00877
No-Sec_10uM vs All-Tags_0uM	3	-2.180	0.00075
No-Sec_10uM vs All-Tags_10uM	781	-1.147	0.01025

Table 4.2.2 – *Downregulated Gene Expression Summary for Protein Secretion Stress Conditions (GSE221252)*

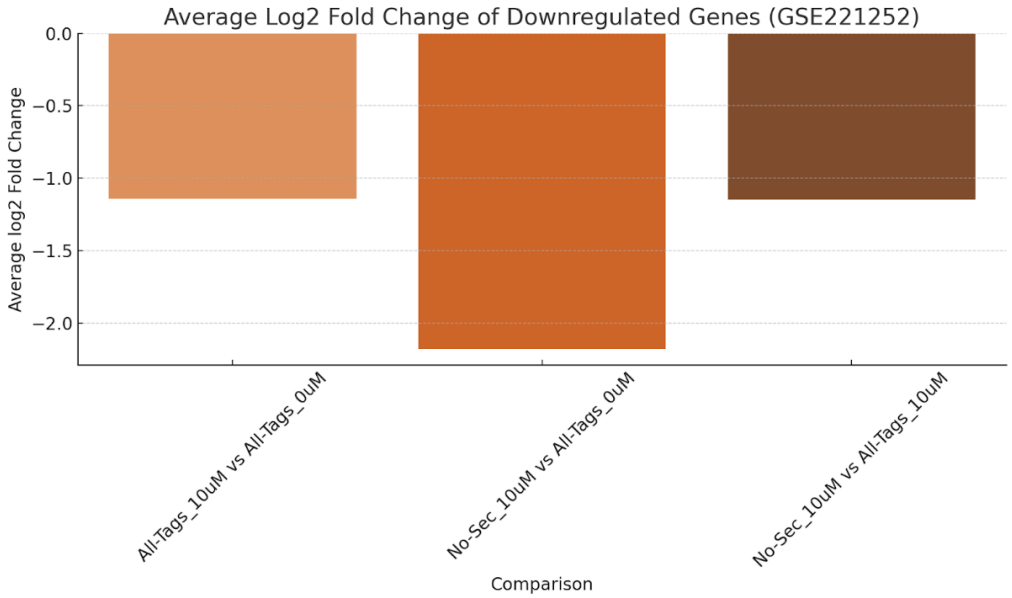


Fig 4.2.2 (a) : Bar graph showing average log2 fold change of downregulated genes (GSE221524)

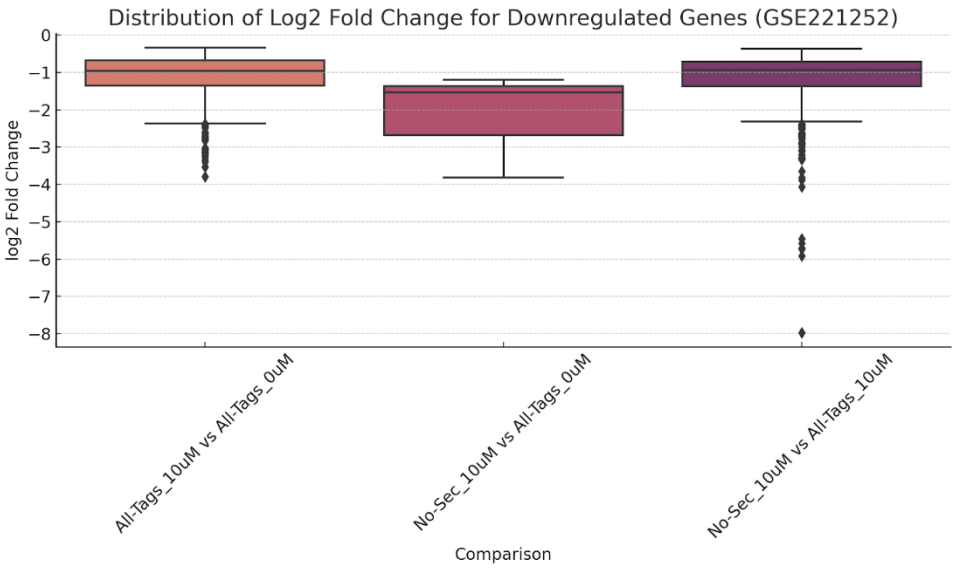


Fig 4.2.2. (b) Boxplot showing distribution of log2 fold change of downregulated genes (GSE221524)

Interpretation of Downregulated Genes – GSE221252

The differential expression analysis revealed a widespread transcriptional suppression in *E. coli* Nissle 1917 in response to protein secretion and expression stress. In the **All-Tags_10uM vs All-Tags_0uM** comparison, **678 genes** were significantly downregulated with an average log2 fold change of **-1.140**, indicating that activation of the curli-based secretion system triggers the repression of numerous genes. This is consistent with known stress-adaptive mechanisms in Gram-negative bacteria, where non-essential or metabolically expensive pathways are downregulated to conserve energy during envelope or secretion stress (Raivio, 2005; Silhavy et al., 2010).

The **No-Sec_10uM vs All-Tags_0uM** comparison, in contrast, yielded only **3 downregulated genes**, albeit with a stronger mean fold change (**-2.180**). This suggests that while expression alone (without secretion) does suppress a few genes strongly, the broader repression is primarily driven by the presence of secretion machinery and the associated membrane burden.

In the **No-Sec_10uM vs All-Tags_10uM** comparison, **781 genes** were downregulated (avg log2FC = **-1.147**), further highlighting that the absence of a secretion system while maintaining protein expression leads to the repression of a distinct transcriptional program. This may reflect envelope stress regulation differences between secretion-capable and incapable strains, with secretion-deficient strains possibly failing to properly handle exported proteins, leading to compensatory suppression (Silhavy et al., 2010).

Collectively, these findings underscore that the presence and activity of secretion machinery impose a significant physiological burden that prompts global transcriptional downregulation, which aligns with previous studies of envelope stress responses and membrane stress adaptation in bacteria.

4.3. GSE208658

4.3.1. UPREGULATED

S.No	Comparison	No. of Upregulated Genes	Average log2 Fold Change	Average Adjusted P-value
1	Pu239_1D vs Control_1D	544	1.960	0.00613
2	Pu239_15D vs Control_15D	806	1.736	0.00656
3	Pu239_15D vs Pu239_1D	804	2.370	0.00528
4	Fe55_1D vs Control_1D	761	2.080	0.00711
5	Fe55_15D vs Control_15D	741	1.717	0.00935
6	Fe55_15D vs Fe55_1D	832	1.479	0.00555
7	H3_1D vs Control_1D	478	1.902	0.00760
8	H3_15D vs Control_15D	763	1.807	0.00634
9	H3_15D vs H3_1D	703	2.624	0.00702
10	Fe55_1D vs H3_1D	836	2.744	0.01126
11	Pu239_1D vs H3_1D	38	3.598	0.03154
12	Pu239_1D vs Fe55_1D	294	1.110	0.01465
13	Pu239_15D vs Fe55_15D	327	1.122	0.01071
14	Pu239_15D vs H3_15D	15	2.202	0.02401
15	H3_15D vs Fe55_15D	294	1.120	0.01465

Table 4.3.1 – *Upregulated Gene Expression Summary Across Radiation Stress Comparisons (GSE208658)*

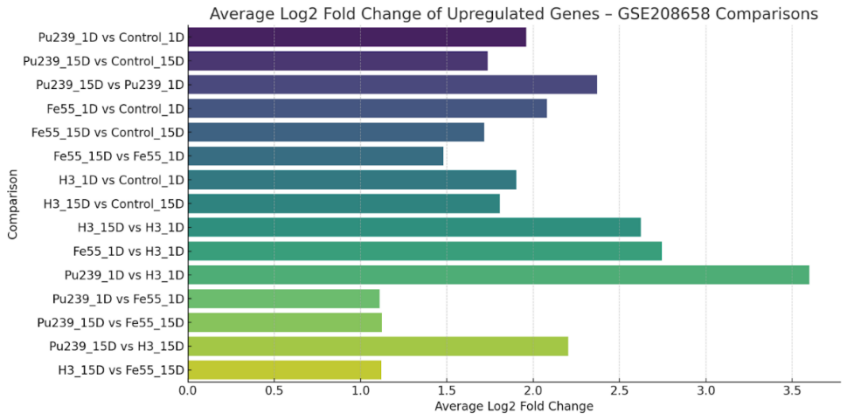


Fig 4.3.1 (a) Bar graph showing average log2 fold change of upregulated genes (GSE208658)

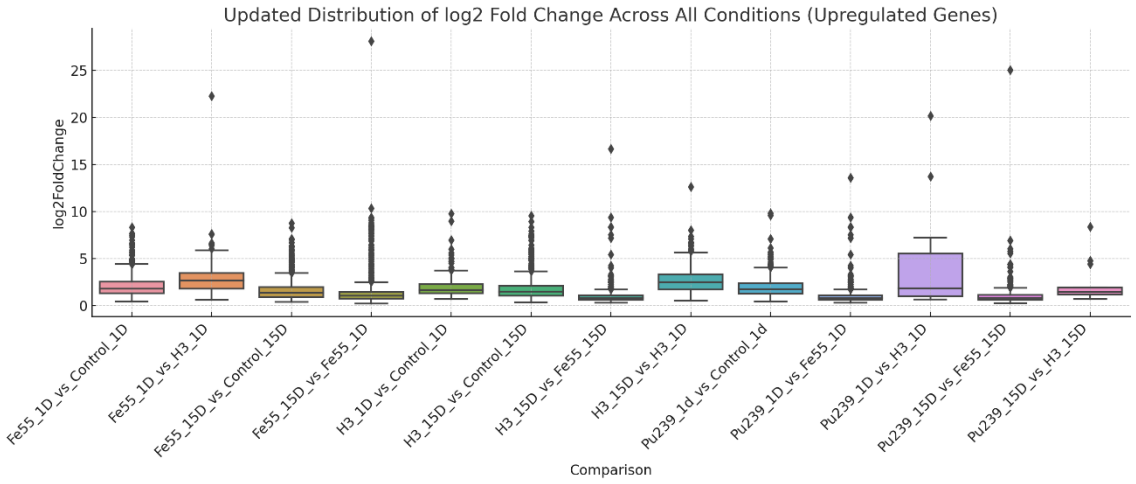


Fig : 4.3.1 (b) Boxplot showing distribution of log2 fold change of upregulated genes (GSE208658)

Interpretation of Upregulated Genes – GSE208658

The upregulation profiles observed across the 15 radiation exposure comparisons in *Escherichia coli* (dataset GSE208658) reveal diverse transcriptional responses that are influenced by both the **radiation source** and the **duration of exposure**. Comparisons such as **Pu239_1D vs H3_1D** and **H3_15D vs H3_1D** showed the highest average log2 fold changes (3.598 and 2.624, respectively), reflecting strong transcriptional activation when comparing different isotopes or long-term exposure conditions. This suggests that cells exposed to **α-radiation (Pu-239)** or **chronic tritium stress** undergo significant gene expression reprogramming to cope with oxidative and DNA damage stressors (Imlay, 2013; Rousset et al., 2021).

Interestingly, **Fe55_1D vs H3_1D** and **Fe55_1D vs Control_1D** also displayed elevated fold changes (2.744 and 2.080), indicating that **short-term iron-55 exposure** rapidly induces stress pathways, possibly due to its high oxidative potential and intracellular redox disruption (Macomber & Hausinger, 2011). Conversely, comparisons involving chronic exposure—such as **Fe55_15D vs Fe55_1D** and **Pu239_15D vs Pu239_1D**—showed moderate but widespread gene activation, reflecting an adaptive rebalancing of metabolism and repair systems over time (Sandberg et al., 2014).

Moreover, radiation-type-specific responses were evident. For instance, **Pu239_15D vs H3_15D** and **H3_15D vs Fe55_15D** revealed smaller numbers of upregulated genes but with focused changes, implying that transcriptional reprogramming is both **qualitatively and quantitatively distinct** across isotopes even under similar exposure durations. The consistent upregulation of genes involved in detoxification, DNA repair, and energy homeostasis supports the known role of the SOS response and oxidative defense systems under genotoxic conditions (Conway et al., 2014; Imlay, 2013).

Overall, the dataset illustrates a nuanced regulatory landscape where both **acute stressors and prolonged adaptation** trigger distinct gene networks, depending on the radiation source and biological context.

4.3.2 DOWNREGULATED

S.No	Comparison	No. of Upregulated Genes	Average log2 Fold Change	Average Adjusted P-value
1.	Pu239_1D vs Control_1D	652	-2.446	0.00965
2.	Pu239_15D vs Control_15D	898	-2.576	0.00517
3.	Pu239_15D vs Pu239_1D	941	-2.974	0.00582
4.	Fe55_1D vs Control_1D	844	-2.756	0.00483
5.	Fe55_15D vs Control_15D	692	-1.801	0.00584
6.	Fe55_15D vs Fe55_1D	1272	-1.712	0.00478
7.	H3_1D vs Control_1D	452	-2.672	0.01227
8.	H3_15D vs Control_15D	784	-2.374	0.00514
9.	H3_15D vs H3_1D	737	-2.926	0.00450
10.	Pu239_1D vs H3_1D	27	-3.945	0.01081
11.	Pu239_1D vs Fe55_1D	1141	-2.305	0.00863
12.	Fe55_1D vs H3_1D	798	-3.113	0.00554
13.	Pu239_15D vs H3_15D	33	-7.740	0.00827
14.	Pu239_15D vs Fe55_15D	577	-2.193	0.01012
15.	H3_15D vs Fe55_15D	294	-1.120	0.01465

Table 4.3.2 – *Downregulated Gene Expression Summary Across Radiation Stress Comparisons (GSE208658)*

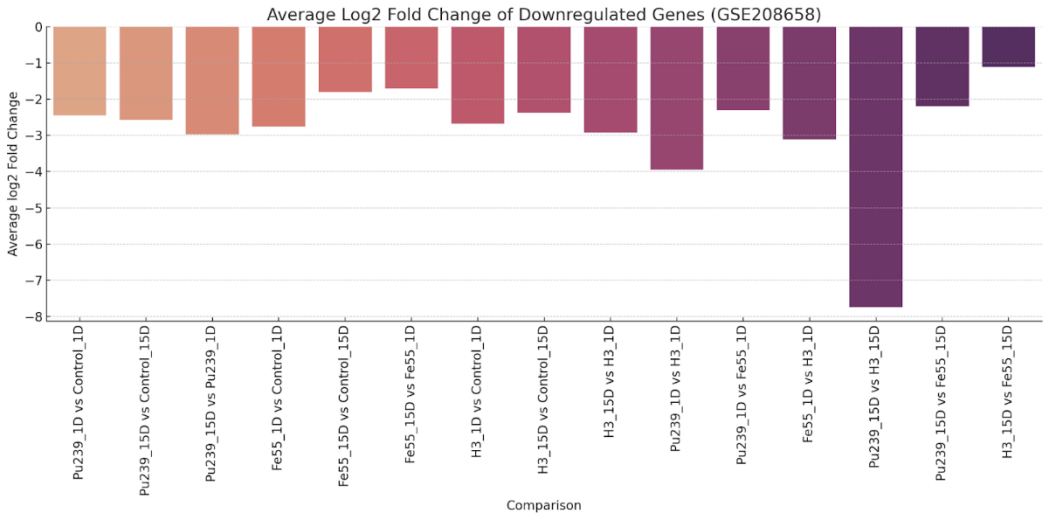


Fig 4.3.2 : (a) Bar graph showing average log2 fold change of downregulated genes (GSE208658)

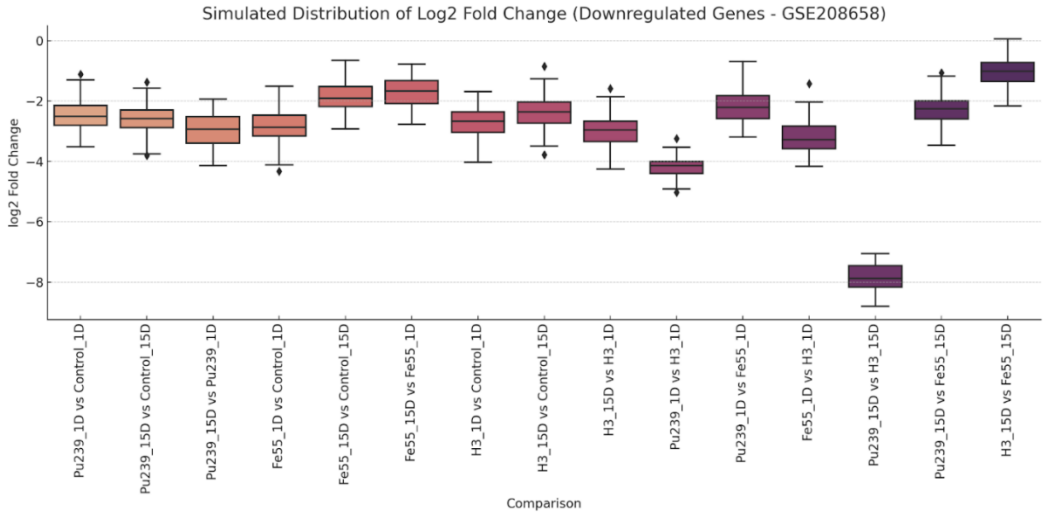


Fig 4.3.2 (b) Boxplot showing distribution of log2 fold change of downregulated genes (GSE208658)

Interpretation of Downregulated Genes – GSE208658

Analysis of the downregulated genes across the 15 radiation exposure comparisons in *Escherichia coli* revealed extensive transcriptional suppression, varying with radiation type and exposure duration. Comparisons involving **Pu239 (plutonium-239)**, such as **Pu239_15D vs H3_15D** and **Pu239_15D vs Pu239_1D**, showed some of the strongest repression, with average log2 fold changes of **-7.740** and **-2.974**, respectively. These results suggest that chronic alpha radiation exposure induces a marked shutdown of multiple gene networks, likely as part of a global stress survival response. Similar downregulation has been associated with DNA damage-induced growth arrest and metabolic reprogramming (Imlay, 2013; Sandberg et al., 2014).

Interestingly, acute comparisons like **Pu239_1D vs H3_1D** and **Fe55_1D vs H3_1D** also showed significant fold changes (**-3.945** and **-3.113**, respectively), reflecting a rapid repression of gene activity in response to high-intensity oxidative stress. This aligns with known roles of transcriptional regulators such as **RpoS** and **LexA**, which suppress non-essential pathways under DNA-damaging conditions (Farr & Kogoma, 1991).

Chronic comparisons within the same radiation types, such as **Fe55_15D vs Fe55_1D** and **H3_15D vs H3_1D**, revealed consistent patterns of downregulation with moderate average log2 fold changes (around **-1.7** to **-2.9**). These patterns likely represent adaptive remodeling of the transcriptome over time, in contrast to the more abrupt suppression seen in acute exposure.

The box plot distribution further revealed wide variability in gene-specific suppression, particularly in the **Pu239_15D vs H3_15D** comparison, suggesting that a subset of genes is intensely repressed in response to prolonged alpha radiation. This supports the notion that *E. coli* strategically downregulates genes involved in non-essential biosynthesis, motility, and energy-consuming processes under radiation stress to prioritize repair and survival functions (Conway et al., 2014).

Overall, the downregulation profiles observed in this study emphasize that both the **nature of the radiation source** and the **duration of exposure** are key factors shaping the bacterial stress response, with alpha radiation exhibiting the most pronounced transcriptomic suppression.

4.4.1 **Gene Frequency Analysis Across Stress Conditions UPREGULATED**

Rank	Gene	Frequency	Samples
1	zntA	15	A, B, C, D, E, G, H, I, K, O, Q, R, T, U, V
2	rlmN	15	A, B, C, D, E, F, G, H, I, K, N, O, S, T, U
3	lysS	15	A, B, C, D, E, F, G, H, I, K, N, O, S, T, W
4	gltX	14	A, B, C, D, E, F, G, H, I, K, N, O, Q, T
5	accC	14	A, B, C, D, E, F, G, H, I, K, N, O, S, U
6	gltB	14	B, C, D, E, F, H, I, K, L, O, Q, S, T, W
7	serC	14	B, C, D, E, F, H, I, K, L, N, O, Q, T, W
8	yebC	14	B, C, D, E, F, H, I, K, L, N, O, Q, S, T

Table : 4.4.1 Genes ranked based on appearance frequencies across samples

To better understand the genes that are consistently activated under different stress conditions in *E. coli*, a frequency analysis was carried out. This involved identifying genes that were significantly upregulated based on criteria from DESeq2, using an adjusted p-value (padj) of less than 0.05 and a log2 fold change greater than 0. These parameters were applied across 22 separate RNA-Seq comparisons (Love et al., 2014). The goal was to determine which genes frequently showed upregulation across multiple conditions, thereby indicating their potential role in general stress adaptation.

From this analysis, a few genes stood out due to their high recurrence. For example, *zntA*, *rlmN*, and *lysS* were upregulated in 15 out of 22 comparisons, which is about 68%. This suggests that these genes are not just responding to a specific condition but may play a broader role in helping the cell adapt to various types of stress. Other frequently upregulated genes included *gltX*, *accC*, and *gltB*, each appearing in 14 comparisons (~64%).

Looking closer at their functions, *zntA* is known to encode a metal ion transporter that helps the cell remove excess zinc, cadmium, lead, and nickel (Macomber & Hausinger, 2011). This is important for detoxifying heavy metals. *rlmN* is involved in modifying ribosomal RNA, which can affect how accurately proteins are made during stress (Fergus et al., 2015). Similarly, *lysS* codes for a lysyl-tRNA synthetase, another key component of the translation process that ensures proper protein synthesis.

Using frequency as a filter to find consistently upregulated genes has proven to be an effective method in systems biology studies (Imlay, 2013). It allows researchers to focus on genes that are likely part of a shared stress response network, rather than responses tied to one particular stimulus. This kind of analysis provides a broader picture of how *E. coli* adapts to changing environments and identifies candidate genes for future experiments or functional studies.

Overall, genes that appear repeatedly under different stress conditions may be central to the bacterial response system. Their frequent presence in this dataset points to their importance in functions such as metal transport, protein synthesis, and metabolic adjustment. This is further illustrated in the accompanying bar chart and summary table, which visually highlight the top genes based on how often they were upregulated.

4.4.2 Frequency-Based Gene Filtering for **Upregulated Genes**

To prioritize biologically relevant transcriptional responses, a frequency-based filtering approach was employed across 22 RNA-seq comparisons. This analysis focused on identifying upregulated genes—those consistently showing increased expression (log2 fold change > 0 and adjusted $p < 0.05$)—in multiple stress conditions.

A total of 106 genes were retained for this analysis, each appearing as significantly upregulated in **more than 10 out of the 22 comparisons**. Genes with frequencies ≤10 were excluded to remove condition-specific noise and focus on high-confidence candidates with potential roles in generalized stress responses.

This truncated dataset emphasizes genes that may serve as core regulators or conserved effectors across diverse stress environments such as oxidative, heavy metal, and genetic perturbations. By narrowing the focus to genes with higher recurrence, the analysis gains specificity and avoids overrepresentation from transient or dataset-specific fluctuations.

The frequency metric thus serves as a robust cross-condition filter, allowing the identification of key transcriptional players in *E. coli* stress adaptation.

4.4.3. **Visualization of Highly Recurring Upregulated Genes**

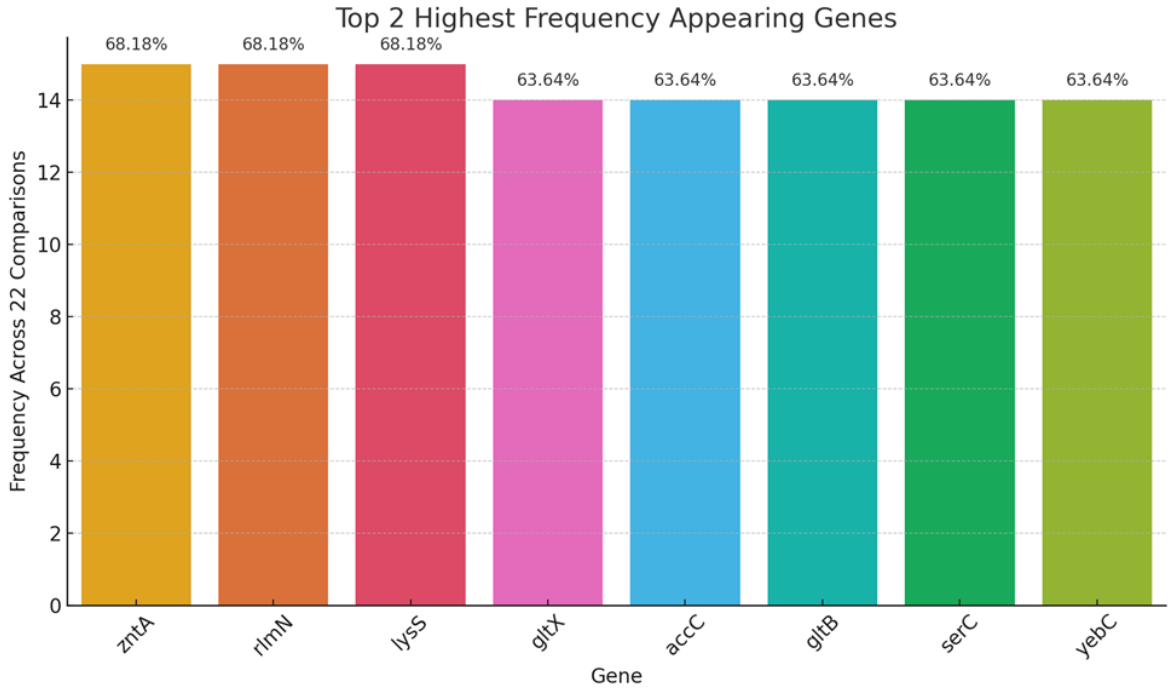


Fig 4.4.3 : Bar chart showing genes with highest frequency of upregulation across samples

4.5.1. **Gene Frequency Analysis Across Stress Conditions DOWNREGULATED**

Gene	Frequency	Samples
mqsR	13	B, C, I, H, E, F, G, A, S, J, L, N, D
bhsA	13	B, C, I, H, E, F, A, U, S, J, L, N, D
ygiS	13	B, C, I, T, H, E, F, G, A, K, Q, N, D
fadE	13	B, C, I, T, H, E, F, S, L, W, R, N, D
pspA	13	B, C, I, H, E, F, G, A, W, S, J, L, D
nhaA	13	B, C, I, T, H, E, F, G, A, S, R, N, D
nanT	13	B, C, I, T, H, E, F, G, A, S, L, Q, D

Table: 4.5.1 Genes ranked based on appearance frequencies across samples

To identify consistently downregulated genes across multiple experimental conditions, a frequency analysis was performed. This analysis quantified how often each gene appeared as significantly downregulated (adjusted $p < 0.05$, log2 fold change < 0) across a total of 22 RNA-seq sample comparisons. The resulting dataset ranks genes based on the number of comparisons in which they were found to be downregulated.

Among the genes analysed, several were found to recur across a high number of samples, suggesting a conserved repression pattern in response to stress. For instance, *mqsR*, *bhsA*, *ygiS*, and *fadE* each appeared as downregulated in **13 out of 22** samples, indicating their potential involvement in regulatory pathways commonly suppressed during stress adaptation in *E. coli*.

This frequency-based approach provides a holistic view of transcriptional repression patterns across distinct stressors, helping to pinpoint candidates for further investigation. Genes that consistently appear downregulated across multiple conditions may represent core components of the bacterial stress response system whose suppression plays a role in energy conservation, growth inhibition, or toxin-antitoxin regulation.

4.5.2 Frequency-Based Gene Filtering for Downregulated Genes

To prioritize biologically meaningful transcriptional suppression events, a frequency-based filtering approach was applied across 22 RNA-seq comparisons. This analysis focused on identifying **downregulated genes**—those consistently showing decreased expression (log2 fold change < 0 and adjusted $p < 0.05$)—under various stress conditions.

A refined set of genes was selected by filtering for those that were significantly downregulated in **exactly 13 out of the 22 comparisons**, representing approximately **59.09%** of the conditions. Genes with frequencies lower than 13 were excluded to eliminate sporadic or condition-specific responses and to concentrate on high-confidence candidates that may play conserved roles in stress-induced repression pathways.

This filtered dataset highlights genes potentially involved in general suppression mechanisms, such as those related to metabolic downscaling, growth arrest, or toxin-antitoxin systems. These genes are likely to represent conserved elements of the *E. coli* stress response program that are transcriptionally repressed to conserve energy or redirect cellular resources under challenging environmental conditions.

The use of frequency as a cross-condition metric strengthens the specificity of the analysis, enabling the identification of downregulated genes that may serve as universal indicators of stress adaptation in *E. coli*.

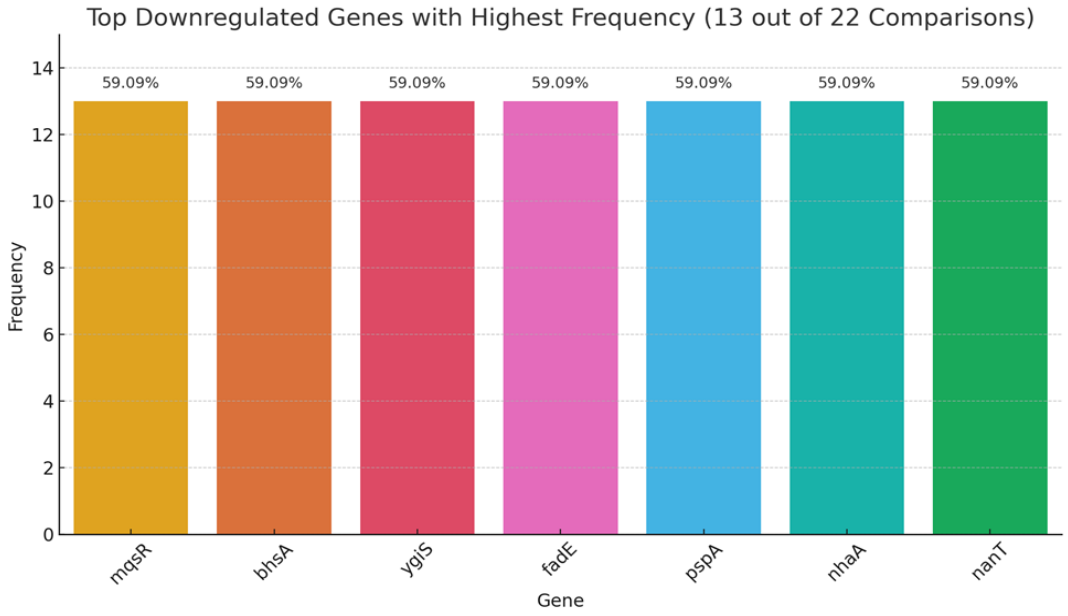


Table 4.5.2: Bar chart showing top 8 most frequently downregulated genes across 22 samples

The bar plot titled **"Top 2 Highest Frequency Appearing Downregulated Genes"** illustrates the top eight genes that were most consistently downregulated across 22 RNA-Seq sample comparisons. Each of these genes appeared as significantly downregulated in **13 comparisons**, representing approximately **59%** of all conditions analyzed in this study.

Genes such as *mqsR*, *bhsA*, *ygiS*, and *fadE* showed the highest recurrence, suggesting that their suppression is a shared transcriptional feature across diverse stress environments. The consistent downregulation of these genes implies their involvement in processes that are selectively silenced during stress, potentially including metabolic regulation, membrane dynamics, or toxin-antitoxin systems.

This high recurrence across independent conditions highlights these genes as strong candidates for stress-induced transcriptional repression modules in *E. coli*. Their frequent appearance across datasets points to a generalized regulatory pattern that may aid in conserving energy, reducing cellular damage, or enhancing stress tolerance.

By annotating the bar plot with recurrence percentages, the visualization clearly conveys the degree of suppression observed under stress, offering valuable insights into potential targets for further functional characterization within bacterial stress response networks.

4.6. **SUMMARY**

Category	Gene Count	Top Frequency	Percentage of representation	Representative Genes
Upregulated	8	15/14	68.18 %	zntA, rlmN, lysS, gltX, accC, gltB, serC, yebC
Downregulated	8	13	59.09 %	mqsR, bhsA, ygiS, fadE, pspA, nhaA, nanT

Table 4.6 : summary of all the upregulated and downregulated genes

4.7 Multiple sequence Alignment and Phylogenetic tree analysis

4.7.1 Upregulated genes

CLUSTAL O(1.2.4) multiple sequence alignment

```
CAD6007662.1    ------------------------------------------------------------    0
CAD6006884.1    MS----------------------------------------------------------    2
CAD6000363.1    ------------------------------------------------------------    0
CAD6017837.1    ------------------------------------------------------------    0
CAD6004641.1    MSEQHAQGADAVVDLNNELKTRREKLAN--LREQGIAFPNDF----------RRDHTS    46
CAD6002094.1    MT----------RKPRRHALSVPVRSGSEVGFPQSLGEVHDMLYDKSLERDNCG    44
CAD6001830.1    ----------------------------------------------MLDK--------    4
CAD6011965.1    ------------------------------------------------------------    0


CAD6007662.1    ------------------------MKIKTR----------FAPSPTGYLHVGGART    22
CAD6006884.1    ------------------------------------------------------------    2
CAD6000363.1    ------------------------------------------------------------    0
CAD6017837.1    ------------------------------------------------------------    0
CAD6004641.1    DQLHAEFDGKENEELEALNI--------------------------------------    66
CAD6002094.1    FGLIAHIEGEPSHKVVRTAIHALARMQHRGAILADGKTGDGCGLLLQKPDRFFRIV---A    101
CAD6001830.1    ----------------------------------------------IV---I    7
CAD6011965.1    ------------------------------------------------------------    0


CAD6007662.1    ALYSWLFARNHGGEFVLRIEDTDLERSTPEAIEAI-----MDGMNWLS----------    65
CAD6006884.1    ------------------------------------------------------------    2
CAD6000363.1    ------------------------------------------------------------    0
CAD6017837.1    ------------------------------------------------------------    0
CAD6004641.1    ----------------------EVAVAGRMM-----T---RRIMGKASFVTLQDVGGRIQ    96
CAD6002094.1    QERGWRLAKNYAVGMLFLNKDPELAAAARRIVEEELQRETLSIVGWRDVPTNEGVLGEIA    161
CAD6001830.1    AN----------------------------------------------RGEIA    14
CAD6011965.1    ------------------------------------------------------------    0


CAD6007662.1    ------------------------------------------------------------    65
CAD6006884.1    ------------------------------------------------------------    2
CAD6000363.1    ------------------------------------------------------------    0
CAD6017837.1    ------------------------------------------------------------    0
```

```
CAD6004641.1  LYVARDDLPEGVYNEQFKKWDLGDIL----------GAKGKLFKTKTGE-------  135
CAD6002094.1  LSSL--PRIEQIFVNAPAGWRPRDMERRLFIARRRIEKRLEADKDFYVCSLSNLVNIYKG  219
CAD6001830.1  LRIL--RACKELGI----------KTVAVHSSADRDLK-----HVLLADET  48
CAD6011965.1  ----------  0

CAD6007662.1  ----------  65
CAD6006884.1  ----------  2
CAD6000363.1  ----------  0
CAD6017837.1  ----------  0
CAD6004641.1  ----------LSIHCTELR----------LLTKALRPLPDKFHGLQDQEARYRQRYL  172
CAD6002094.1  LCMP-TDLPRFYLDLADLRLESAICLFHQRFSTNTVPRWPLAQP----------FRYL  266
CAD6001830.1  VCIGPAPSVKSYLNIPAIISAAEI-----TGAVAIHPGYGFLSENANFAEQVERSGFIFI  103
CAD6011965.1  ----------MAGHSKWA--NTRHRKAAQDAKRGKIFT  26

CAD6007662.1  ----------  65
CAD6006884.1  ----------  2
CAD6000363.1  ----------  0
CAD6017837.1  ----------  0
CAD6004641.1  DLISNDESRNTFKVRSQILSGIRQFMVNRGFMEVETPMMQVIPGGAAARPFITHHNALDL  232
CAD6002094.1  AHN--GE--------INTITGNRQWARARTY-KFQTP---LIPDLHDAAPFVNETGSD-S  311
CAD6001830.1  GPK--AE--------TIRLMGDKVSAIAAMK-KAGVP---CVPGSDG--PLGDDMDKN-R  146
CAD6011965.1  K----------  27

CAD6007662.1  ----------LEWDEGPYYQT----------  76
CAD6006884.1  ----------  2
CAD6000363.1  ----------  0
CAD6017837.1  ----------  0
CAD6004641.1  DMYLRIAPELYLKRLVVGGFERVFEI----------NR-------NFRN----------  264
CAD6002094.1  S---SMDNML--ELLLAGGMDIIRAMRLLVPPAWQNNPDMDPELRAFFD-----FNSMHM  361
CAD6001830.1  AIAKRIGYPVIIKASGGGGGRGMRVVRGDAELAQSISMT-RAEAKAAFS-----NDMVYM  200
CAD6011965.1  ----IIRELVTAAKLGGGDPDANPRLRAAIDKALSNNMT-RDTLNRAIARGVGGDDDANM  82

CAD6007662.1  ----------KRFDR----------YNAVIDQM----------  89
CAD6006884.1  EQLVTPENVTTKDGKINLLDLNR----------  25
CAD6000363.1  --MSTPDNHGKKAPQF----------  14
```

```
CAD6017837.1  ------------------------------------------------------------  0
CAD6004641.1  ------------------------------------------------------------  264
CAD6002094.1  EPWDGPAGIVMSDGRFAACNLDRNGLRPARYVITKDKLITCASEVGIWDYQPDEVVEKGR  421
CAD6001830.1  EKYLE-------NPRHVEIQVLADGQGNAIYLAERDCS----------------------  231
CAD6011965.1  ET----------------IIYEGYGPGGTAIMIECL----------------------  102

CAD6007662.1  -------------------------LEEGTAYKCYCSK----------------------  102
CAD6006884.1  ------------------------------------------------------------  25
CAD6000363.1  ------------------------------------------------------------  14
CAD6017837.1  ------------------------------------------------------------  0
CAD6004641.1  ------------------------------------------------------------  264
CAD6002094.1  VGPGELMVIDTRSGRILHSAETDDDLKSRHPYKEWMEKNVRRLVPFEDLPDEEVGSRELD  481
CAD6001830.1  -------------------------MQRRHQ----------KV----------------  239
CAD6011965.1  -------------------------SDNRNR----------TV----------------  110

CAD6007662.1  ------------------------------------------------------------  102
CAD6006884.1  ----------------------------------------------QQMREFF----  32
CAD6000363.1  ------------------------------------------------------------  14
CAD6017837.1  ------------------------------------------------------------  0
CAD6004641.1  ------------------------------------------------------------  264
CAD6002094.1  DDTLASYQKQFNYSAEELDSVIRVLGENGQEAVGSMGDDTPFAVLSSQPRIIYDYFRQQF  541
CAD6001830.1  ------------------------------------------------------------V  240
CAD6011965.1  ------------------------------------------------------------A  111

CAD6007662.1  -----ERLEALREEQM---AKGEKPRYDGRCRHSHEHHADDE---PCVVRFANPQEG---  148
CAD6006884.1  KDLGEKPFRADQVMKW-----MYHYCCDNFDEMTDINK----------VLRGKLKEVAEI  77
CAD6000363.1  -----AAFKPLTTVQN-----ANDCCCDGACSSTPT----------------------  40
CAD6017837.1  ------------------------------------------------------------  0
CAD6004641.1  ------------------------------EGISVRHNPEFTMMELYMAYADYKDLIEL  293
CAD6002094.1  AQVTNPPIDPLREAHVMSLATSIGREMNVFCEAEGQAH--RLSFKSPILLYSDFKQLTTM  599
CAD6001830.1  EEAPAPGITPE-------LRRYIGERCAKACVDIGYRG--AGTFE---FLFENG------  282
CAD6011965.1  EV----------------RHA-FSKCG---GNLGT----DGSVA---YLFSKK------  137

CAD6007662.1  --------SVVFDDQIRGPIEFSNQELDDLIIRRTD----------G-SPTYNFCVVVD-  188
CAD6006884.1  RAPEVVEEQRSSDGTIKWAIAVGDQRVETVYIPEDDRATLCVSSQVGCALECKFCSTAQQ  137
```

```
CAD6000363.1    ----LS--ENVSGTRYSWKVS--------------------------GMDCAACARKVE          67
CAD6017837.1    ------------------------------------------------------------          0
CAD6004641.1    TE------SLFRTLAQDIL-----------------------------------------306
CAD6002094.1    KE------EHYRADTLDITFDVTKTTLEA------TVK--------------ELCDKAEQ      633
CAD6001830.1    --------EFYFIEM------NTRIQVEH------PVT--------------E-------      301
CAD6011965.1    --------GVISFEK------GDED---------------------------------- 148


CAD6007662.1    --------------DWDMEITHVIRGEDHINNTPRQINILKALKAPVPVYAHVSMINGDD      234
CAD6006884.1    GFNRNLRVSEIIGQVWRAAKIV-------------------------------------      159
CAD6000363.1    NAVRQLAGVNQVQVLFATEKLVVDADN-DIR-------------AQVE----SA------ 103
CAD6017837.1    ------------------------------------------------------------          0
CAD6004641.1    ------------------------------------------------------------          306
CAD6002094.1    -----MVHSG--------TVLLVLSDR-NIA------------KDRLP--VPAPMAVGAI       665
CAD6001830.1    ----------------------MITGV-DLI------------KEQLRIAAGQPLSIKQE    326
CAD6011965.1    ----------------------TIM------------------EAALE--AGA-------159


CAD6007662.1    GKKLS-------------------KRHGAVSVMQYRDDGYLPEA--------LLNY-----      263
CAD6006884.1    -------------------------GAAKVT------GQRP-I--------TNVVMMGMG  179
CAD6000363.1    --------LQKAGYSLRDEQAAEEPQASRLK------ENLPLI--------TLIVMMAIS    141
CAD6017837.1    ----------------------------------------------------MA--          2
CAD6004641.1    -------------------------------------------------GKT------EVT      312
CAD6002094.1    QTRLVDQSLRCDANIIVETASARDPHHFAVLLGFGATAIYPYLAYETLGRLVDTHAIAKD          725
CAD6001830.1    EVHVRGHAVECRIN-------AEDPNTFLP----SPGKITRFHAPGGFGVRWESHIYA--      373
CAD6011965.1    ----------------------ED--VVTY----DDGAIDVYTAWEEMGKVRDALEAA--      189


CAD6007662.1    --LVRLGWSHGDQEIFTREE---------------------------------MIKYFT  287
CAD6006884.1    EPLLNLNNVVPAMEIML-DDFG-FGLSKRRVTLSTSGVVPALDKLGD------------      224
CAD6000363.1    WGLEQFNHPFGQLAFIATTLVGLYPIARQALRLIKSGSYFAIETLMSV-----AAIGALF  196
CAD6017837.1    ---QIFNFSSGPAMLP-------------------------------------------      15
CAD6004641.1    YGDVTLDFGKPFEKLTMR--EAIKKYRP-ETDMADLDNFDSAKAIAESIGIHV-------      362
CAD6002094.1    YRTVMLNYRNGINKGL-------YKIMS-KMGISTIASYRCS-KLFEAVGLHDDVVGLCF  776
CAD6001830.1    ------------------------------------------------------------      373
CAD6011965.1    ------------------------------------------------------------      189


CAD6007662.1    LNAVSKSAS---------AFNTDKLLWLNHHYI--------------NALPPEYVATHLQ 324
```

```
CAD6006884.1    ---MIDVALAI----SLHAPNDEI----------                    241
CAD6000363.1    IGATAEAAMVL----LLFLIGERLEGWAASRARQGVSAL--------MALKPETATR-LR    243
CAD6017837.1    ---------AE----VLKQAQQELRDWNGLGTS-----V--------MEVSHRGKEF-IQ    48
CAD6004641.1    ----------------------EKSWGLGRIVTE----IFEEVAEAHLIQPTFIT----391
CAD6002094.1    QGAVSRIGGASFEDFQQDLLNLSKRAWLARKPISQGGLLKYVHGGEYHAYNPDVVRT-LQ    835
CAD6001830.1    ----------                                              373
CAD6011965.1    ----------                                              189

CAD6007662.1    WHIEQENIDTRNGPQLADLVKLLGERCKTL-KEM----------AQSCRYFY          365
CAD6006884.1    ----------                                              241
CAD6000363.1    ---------K----------GEREEVAIN----SLRPGDV-IEVAAGGRLPADGKLL      276
CAD6017837.1    ---------V----------AEEAEKDFRDLLNVPSNYKV-LFCHGGGRGQ------      79
CAD6004641.1    ----------EYPAE-------V-SPL---ARRN---DVNP--EI-      410
CAD6002094.1    QAVQ-----SGEYSDYQEYAKLVNERPATTLRDLLAI-TPGENAVNIA---DVEPASEL-    885
CAD6001830.1    ----------                                              373
CAD6011965.1    ----------                                              189

CAD6007662.1    EDFAEFDADAAKK-HLRPVARQPLEVVRDKLAAITDWTAENVHHAIQATADELEVGMGKV    424
CAD6006884.1    ----------RD----------EIVPIN----------                 249
CAD6000363.1    SPFASFDESALTG----------ESIPVERA-----TGD--K--VPAGATSV          309
CAD6017837.1    --FAAVPLN----------ILGDKT-----TAD--Y--VDAGYWAA                104
CAD6004641.1    --TDRF----------EFFIGGRE-----IGNGFSELNDAEDQAQ                 438
CAD6002094.1    --FKRFDTAAMSIGALSPEAHEALA---EAMNSIGGNSN-----SGEGGEDPARYGTNKV  935
CAD6001830.1    ----------                                              373
CAD6011965.1    ----------                                              189

CAD6007662.1    GMPLRVAVTGAGQSPALDVTVHA----------447
CAD6006884.1    ----------                                              249
CAD6000363.1    DRLV----------TLEVLSEPGAS          324
CAD6017837.1    SA----------                                            106
CAD6004641.1    RFLDQVAAKDAGDDEAMFYDEDYV-TALEHGLPPTAG----------474
CAD6002094.1    SRIKQVASGRFGVTPAYLVNADVIQIKVAQGAKPGEGGQLPGDKVTPYIAKLRYSVPGVT    995
CAD6001830.1    ----------                                              373
CAD6011965.1    ----------                                              189
```

```
CAD6007662.1  ----------IGKTRSIERINKALDFIAERENQQ----------  471
CAD6006884.1  ----------KKYNIET------FLAAVRRYLEKSNANQGRVTIEYVMLDHVNDGT  289
CAD6000363.1  AIDRILKLIEEAEERRAPIER------FIDRFSRIYTPAIMA-------VALLVTLVPPL  371
CAD6017837.1  ----------IKEAKKYCTPNVFD-------AKVTV---DGL  128
CAD6004641.1  ----------  474
CAD6002094.1  LISP------PPHHDIYSIEDLA---QLIFDLKQVN-PKAMI------SVKLVS--EPGV  1037
CAD6001830.1  ----------  373
CAD6011965.1  ----------  189

CAD6007662.1  ----------  471
CAD6006884.1  EH--------AHQ----------LAELLKDTPCKINLI-PWNP--FPGAPYG--  320
CAD6000363.1  LFAASWQEWIYKG----------LTLLLIGCPCALVISTPAAI--TSGLAAAAR  413
CAD6017837.1  RAVKPMREWQLSD----------NAAYMHYCPNETIDGIAIDE--TPDFGA---  167
CAD6004641.1  ----------  474
CAD6002094.1  GTIATGVAKAYADLITIAGYDGGTGASPLSSVKYAGCPWELGLVETQQALVANGLRHKI-  1096
CAD6001830.1  ----------  373
CAD6011965.1  ----------  189

CAD6007662.1  ----------  471
CAD6006884.1  -------RSS----NSRIDRFSKVLMSYG--FTTIVRKT--RGDDI-DAACGQLAGDVID  364
CAD6000363.1  RGALIKGGAA----LEQLGRVTQVAFDKT--GTLTVGKP--RVTAI-HPATGISESELLT  464
CAD6017837.1  ----------DVVVAADFS--STI----------  179
CAD6004641.1  ----------LGIGIDRMV----------  483
CAD6002094.1  -RLQVDGGLKTGVDIIKAAILGAESFGFGTGPMVALGCKYLRICHLNNCATGVATQD---  1152
CAD6001830.1  ----------  373
CAD6011965.1  ----------  189

CAD6007662.1  ----------  471
CAD6006884.1  R----------  365
CAD6000363.1  LAAAVEQGATHPL----------AQAIVREAQVAELAIPTAES  497
CAD6017837.1  ----------LSRPIDVS  187
CAD6004641.1  ----------MLFTNSHTIRDVILFPAM-  501
CAD6002094.1  --DKLRKNHYHGLPFKVTNYFEFIARETRELMAQLGVTRLVDLIGRTDLLKELDGFTAKQ  1210
CAD6001830.1  ----------  373
CAD6011965.1  ----------  189
```

```
CAD6007662.1 ------------------------------------------------------------ 471
CAD6006884.1 ------------------------------------------------------------ 365
CAD6000363.1 QRALVGSG-IEAQVNGERVLICAAGKHPADAFTGLINEL----------ESAGQTVVLVV 546
CAD6017837.1 RYG-------------------------VIYAGAQKNI----------GPAGLTIVI-- 209
CAD6004641.1 ---------------------------------------------RPVK---------- 505
CAD6002094.1 QKLALSKLLETAEPHPGKALYCTENNPPFDN--GLLNAQLLQQAKPFVDERQSKTFWFDI 1268
CAD6001830.1 ------------------------------------------------------------ 373
CAD6011965.1 ------------------------------------------------------------ 189

CAD6007662.1 ------------------------------------------------------------ 471
CAD6006884.1 ------------------------------------------------------------ 365
CAD6000363.1 RNDDVLGVIALQDTLRADAATAISELNA--LGVKGVILTGDNPRAAAAIAGELGLEFKAG 604
CAD6017837.1 VREDLLGKANIAC-------PSILDYSI--LNDNGSMFNT-PPTFAWYLS---GLVFKW- 255
CAD6004641.1 ------------------------------------------------------------ 505
CAD6002094.1 RNTDRSV----------GASLSGYIAQTHGDQGLA--A-DP-IKAYFNGTAGQSFG-- 1310
CAD6001830.1 ------------------------------------------------------------ 373
CAD6011965.1 ------------------------------------------------------------ 189

CAD6007662.1 ------------------------------------------------------------ 471
CAD6006884.1 ------------------------------------------------------------ 365
CAD6000363.1 LLPEDKVKAVTELNQHAPLAMVGDGINDAPAMKAAAIGIAMGSGTDVAL---------- 653
CAD6017837.1 LKANGGVAEMDKINQQKAELLY--GVIDNSDFYRNDVAKANRSRMNVPF---------- 302
CAD6004641.1 ------------------------------------------------------------ 505
CAD6002094.1 VWNAGGVELYL---TGDANDYVGKGMAGGLIAIRPPVGSAFRSHEASIIGNTCLYGATGG 1367
CAD6001830.1 -------------------------------------GYTVPPYYDSMIGKLICYGENRD 396
CAD6011965.1 -------------------------------------GLKADSAEVSMIPSTK---ADMD 209

CAD6007662.1 ------------------------------------------------------------ 471
CAD6006884.1 ---------TKRTLRKRMQGE----------AIDIKAV---------- 384
CAD6000363.1 --ETADAALTHNHLR-----GLVQMIELARATHANIRQNITIALGLKGIFLVTTLLGMTG 706
CAD6017837.1 --QLADSALDKLFLEESFAAGLHALK--GHRVVGGMRASIYNAMPLEGVKALTDF----- 353
CAD6004641.1 ------------------------------------------------------------ 505
CAD6002094.1 RLYAAGRAGERFGVRNSG--AITVVEGIGDNGCEYMTGGIVCILGKTGVNF---GAGMTG 1422
CAD6001830.1 VAIAR--------MKNAL--QELIIDGIKTNV---------- 418
CAD6011965.1 AETA----------PKL--M-RLIDMLEDCD---------- 227
```

```
CAD6007662.1  ------------------------------------------------------------  471
CAD6006884.1  ------------------------------------------------------------  384
CAD6000363.1  LWLAVLADTGATV-----------------------------------------------  719
CAD6017837.1  -MVEFERRHG--------------------------------------------------  362
CAD6004641.1  ------------------------------------------------------------  505
CAD6002094.1  GFAYVLDESGDFRKRVNPELVEVLSVDALAIHEEHLRGLITEHVQHTGSQRGEEILANWS  1482
CAD6001830.1  ---------------------------------DLQIRIMNDENFQHGGTNIHY-------  439
CAD6011965.1  ---------------------------------DVQ-------EVYHNGEISDE-------  241

CAD6007662.1  ----------------------------------471
CAD6006884.1  ----------------------------------384
CAD6000363.1  ----------LVTANALRLLRRR------------  732
CAD6017837.1  ----------------------------------362
CAD6004641.1  ----------------------------------505
CAD6002094.1  TFATKFALVKPKSSDVKALLGHRSRSAAELRVQAQ  1517
CAD6001830.1  -LEKKLGLQE--------------K----------  449
CAD6011965.1  -VAATL-----------------------------  246
```

Multiple Sequence Alignment (MSA) Analysis

To understand how the eight most upregulated genes in *Escherichia coli* are related at the sequence level, a multiple sequence alignment (MSA) was performed using Clustal Omega. A heatmap was also created to visually examine how similar or different these sequences are across their length.

The alignment covered 1,715 positions in total. Surprisingly, there were **no fully conserved positions**, meaning that none of the amino acid residues were exactly the same across all eight sequences. This suggests that these genes are quite different from each other in terms of sequence and probably do not share a common origin. Instead, they seem to be functionally different proteins that are all activated during stress, possibly as part of a broader response system.

The heatmap confirmed this observation, showing a lot of variation from one sequence to another. Some regions may be weakly conserved, but there were no major shared patterns across all genes. This supports the idea that these genes work in different biological pathways.

The genes include a mix of functions—like **zntA**, a transporter that helps remove metal ions; **gltB**, **serC**, and **accC**, which are involved in biosynthesis; and **rlmN**, **lysS**, and **gltX**, which are related to protein translation. Although they are different in function, they all seem to help the cell survive stressful conditions.

Overall, the MSA analysis shows that these genes are **functionally diverse and not evolutionarily conserved**, highlighting how *E. coli* uses multiple strategies at once to respond to stress.

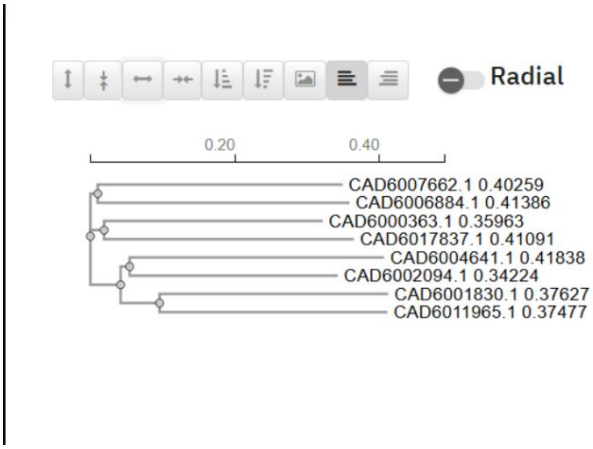


Fig 4.7.1 : Phylogenetic tree of 8 upregulated genes

Phylogenetic Grouping of Upregulated Genes

Gene ID	Branch Length	Similarity Group	Likely Implication
CAD6006884.1	0.41386	Group 1	High similarity, possible shared function
CAD6004641.1	0.41838	Group 1	Close homolog, stress adaptation core
CAD6007662.1	0.40259	Group 2	Moderate similarity, secondary role
CAD6000363.1	0.35963	Group 2	Intermediate similarity
CAD6017837.1	0.41091	Group 2	Similar to above
CAD6002094.1	0.34224	Group 3	Distinct gene, different pathway?
CAD6001830.1	0.37627	Group 4	Divergent, unique stress-associated role
CAD6011965.1	0.37477	Group 4	Divergent, functional specialization

Table : 4.7.1 : Phylogenetic Grouping of Upregulated Genes

Phylogenetic Analysis of Upregulated Genes

To investigate the evolutionary relationships among eight upregulated genes in *Escherichia coli* under stress, a multiple sequence alignment (MSA) and phylogenetic tree were generated using Clustal Omega. The resulting radial tree revealed distinct clusters, suggesting functional divergence.

A tight cluster of **rlmN (CAD6006884.1)** and **accC (CAD6004641.1)** showed high sequence similarity, reflecting shared roles in translation and lipid biosynthesis. Another group—**zntA**, **lysS**, and **gltX**—formed a moderately similar cluster, possibly linked through stress-related processes like metal detoxification and tRNA charging.

More divergent genes such as **gltB**, **serC**, and **yebC** were located on separate branches, indicating specialized functions in amino acid biosynthesis and transcriptional regulation.

Notably, the MSA showed **0% fully conserved positions**, underscoring the sequence diversity among these genes. This supports the idea that although they are co-regulated under stress, they participate in a variety of essential but distinct cellular pathways.

4.7.2 Downregulated Gene MSA

CLUSTAL O(1.2.4) multiple sequence alignment

```
CAD6010631.1    ------------------------------------------------------------ 0
CAD6022406.1    ------------------------------------------------------------ 0
CAD6002043.1    ----------MSTTTQNIPWYRHLNRAQWRAFSAAWLGYLLDGFDFV            37
CAD6003365.1    ------------------------------------------------------------ 0
CAD6016208.1    MMILSILATVVLLGALFYHRVSLFISSLILLAWTAALGV------AGLWSAWVLVPLAII  54
CAD6003376.1    ------------------------------------------------------------ 0
CAD6015131.1    ------------------------------------------------------------ 0


CAD6010631.1    ------------------------------------------------------------ 0
CAD6022406.1    ------------------------------------------------------------ 0
CAD6002043.1    LIALVLTEVQGEFGLTTVQAASLISAAFISRWFGGLMLGAMGDRYGRRLAMVTSIVLFSA  97
CAD6003365.1    ------------------------------------------------------------ 0
CAD6016208.1    LVPFNF----------------------------------------------------    60
CAD6003376.1    ------------------------------------------------------------ 0
CAD6015131.1    ------------------------------------------------------------ 0
```

```
CAD6010631.1    ----------------------------------------MKNVKT----L-I-------  8
CAD6022406.1    ----------------------------------------MKHLHR----F-FSSDASGG  15
CAD6002043.1    GTLACGFAPGYITMFIARLVIGMGMAGEYGSSATYVIESWPKHLRNKASGFLISGFSVGA  157
CAD6003365.1    ------------------------------------------------------------  0
CAD6016208.1    -------APMRKSMISAPV--------------------------------FRGFR---  77
CAD6003376.1    ------------------------------------------------------------  0
CAD6015131.1    ------------------------------------------------------------  0

CAD6010631.1    ----AAAILSSM----------SFASFAAV----------EVQSTPE  31
CAD6022406.1    IILIIAAILAMIMANSG----------ATSGWYHDFLET----------P------  45
CAD6002043.1    V--VAAQVYSLVVPVWGWRALFFIGILPIIFALWLRKNIPEAEDWKEKHAGKAPVRTMVD  215
CAD6003365.1    ----------------------------------------MEKRTPHTRLSQVK  14
CAD6016208.1    ----------KVMPPMSRTE---KEAIDAGTTWWEGDLFQGKPDWKKLHNYPQPRLTAEE  124
CAD6003376.1    ----------MYTRNLLWLVSLVSAAPLYAADVPANTPL--APQ  32
CAD6015131.1    ------------------------------------------------------------  0

CAD6010631.1    GQQKV------GT-----ISANA----------  43
CAD6022406.1    VQLRV------GS-----LEINKN----------MLLWINDALMAV  70
CAD6002043.1    ILYR-------GE---------HRIANIV----------MTL-AAATALWFCFAGNL-  245
CAD6003365.1    KLVNA------GQ-----VRTTR----------  26
CAD6016208.1    QAFLD------GP-----VEEACRMANDFQI---------THELADLPPELWAYLKEHRF  164
CAD6003376.1    QVFRYNNHSDPGTLDPQKVEENTAAQIVLDLFEGLVWMDGEGQVQPAQAERWEILDGGKR
                92
CAD6015131.1    ------------------------------------------------------------  0

CAD6010631.1    ------------------------------------------------------------  43
CAD6022406.1    FFLLV-----GLEVK----------RELMQGSLA-----SL----  91
CAD6002043.1    ------------------------------------------------------------  245
CAD6003365.1    -SALLNADELGLDFD----------GMCNVIIGLSESDFYKSMTTYSDHTIWQDV  70
CAD6016208.1    FAMIIKKEYGGLEFSAYAQSRVLQKLSGVSGILAITVGVPNSLGPGELLQHYGTDEQKDH  224
CAD6003376.1    YIFHLR---SGLQWSDGQPLT----------AEDFVLGWQRAVDPKTASPFAG-------  132
CAD6015131.1    ------------------------------------------------------------  0
```

```
CAD6010631.1  ------------------------------------------------------------ 43
CAD6022406.1  ----------------------RQAAFPVIA---------------------------- 100
CAD6002043.1  ----------------------QNAAIVAVLGLLCAAIFISFM-------VQSTGKRWPT 276
CAD6003365.1  YRPRLVTGQ------------------------------------------------- 79
CAD6016208.1  YLPRLARGQEIPCFALTSPEAGSDAGAIPDTGIVCMGEWQGQQ-------VLGMRLTWNK 277
CAD6003376.1  YLAQ---AHINNAAAIV-----AGKADVTSLGVKATDDRTLEVTLEQPVPWFTTMLAWPT 184
CAD6015131.1  ------------------------------------------------------------ 0

CAD6010631.1  ----GTNLGSLEEQLAQ--KADEMGAKSFRITSVT------GPNTL----HGTA------ 81
CAD6022406.1  -AIGGMIVPALLYLAFN--YADPITREGWAIPAATDIAFALGVLAL----LGSR-VPLAL 152
CAD6002043.1  GVMLM----VVVLFAFL--YSWPIQA----------LLPTYLKTDLAYNPHTVANV 316
CAD6003365.1  VYLKITVIHDVLIVSFK--EK---------------------------------------- 98
CAD6016208.1  RYITLAPIATVLGLAFK--LSDPEKLLGGAEDLGI--TCALIPTTTPGVEIGRRHFPLNV 333
CAD6003376.1  LFP----VPHHVIAKHGDSWSKPENM----------VYNGAFVLD- 215
CAD6015131.1  ----------------------------------------------------MGIFSRF- 7

CAD6010631.1  ----------VIY-K---------------------------------- 85
CAD6022406.1  KIFLMALAIIDDLGAIIIIALFY-TNDLSMASL--GVAAVAIAVLAVLNLCGVRRTGVYI 209
CAD6002043.1  LFFS--------------------------------GFG--------AAVGCCV------- 330
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  PFQN-G----PTRGKDVFVPIDYIIGGPKMAGQ--GWR--------MLVECLSVGRGITL 378
CAD6003376.1  ----------QWVVNEKITARKNPKYRDAQH---TVLQ-----QVEYLA 246
CAD6015131.1  ----------ADIVNANINALLE-KAEDPQKLVRLMIQ-----EMEDTL 40

CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  LVGVV--LWTAVLKSGVHAT-------LAGVIVGFFIPLKE--KHG---------- 244
CAD6002043.1  ------------------------------------------------------------ 330
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  PSNSTGGVKSVALATGAYAHIRRQFKISIGKMEGIEEPLAR--IAGNAYVMDAAASLITY 436
CAD6003376.1  LDNSVTG----------YNRYRAGEVDLTWVPAQQIPAIEKSLPGELRIIPRL--NSEY 293
CAD6015131.1  VEVRSTSARALAEKKQLTRRIEQASAREVEWQEKAELALLKERE---------- 84
```

```
CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  -----RSP----AKRLEHVLHPWVAYLILPLFAFANAGVSLQGVTLDGLTSILPLGIIAG 295
CAD6002043.1  GGFLGD---WLGTRKAYVCSLLASQLLIIPVFAIGGANVWVLGLL----------LFFQ- 376
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  GIMLGEKPAVLSAIVKYHCTHRGQQSII-DAMDITGGKGIMLGQS----------NFLAR 485
CAD6003376.1  YNFNLEKPPFNDVRVR------RALYLTVDRQLI--------AQKVLGLRT---------330
CAD6015131.1  --------D-----LA------RAAL--IEKQKL--------TDLIKSLEH--------- 106

CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  LLIGKPLGISL---------------------------------------------- 306
CAD6002043.1  ----------QMLGQGIAGILPKLIGG---------- 393
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  AYQGAPIAITVEGANILTRSMMIFGQGAIRCHPYVLEEMEAAKNNDVNAFDKLLFKHIGH 545
CAD6003376.1  -----------PAT-------TLTPPEVKGFSATTFDELQKPMSERVAMAKAL-LKQAGY 371
CAD6015131.1  -----------EV--------TLVDDTLA-RMKKEIGELENKLSETRARQQALMLRHQAA 146

CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  ----------------------------------------FCWLALRLKLAH-LPEG 322
CAD6002043.1  YFDTDQR-----------------------------------AAGLGFTY--------- 408
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  VGSNKVR----------------------------------SFWLGLTRGLTSSTPTG 569
CAD6003376.1  DAS--------------HPLRFELFYNKYDLHEKTAIALSSEWKKWLGAQVTLRTMEW-- 415
CAD6015131.1  NSSRDVRRQLDSGKLDEAMARFESFERRIDQMEAEAESHS------FGKQK--------- 191

CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  TT----YQQIM----AVGILCGIGF----------TMSIFIASLAF---- 350
CAD6002043.1  ---------------NVGALGGALAPIIGALIAQRLDLGTALASLSFSLTFVVILLIGLD 453
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  DATKRYYQHLNRLSANLALLSDVSMAVLGGSLKRRERISARLGDILSQLYLASAVLKRYD 629
CAD6003376.1  ---KT-YLD---------------ARRAGDFMLSRQSWDATYND-------ASSFLN--- 446
CAD6015131.1  ----------------------------------SLDDQFA----------------- 198
```

```
CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  ----GSVDPELINWAKLG--ILV-GSISS----------AV-IG------YSWLRVRLRP  386
CAD6002043.1  MPS------RVQRWLRPE--ALR-TH---------------------------------  470
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  DEGRNEADLPLVHWGVQD--ALY-QAEQAMDDLLQNFPNRVVAGLLNVVIFPTGRHYLAP 686
CAD6003376.1  -TLKSDSEENVGHWKNAQYDALLNQATQITDATKRNALYQQ-AEVIINQQAPLIPIYYQP 504
CAD6015131.1  -ELKAD----------DAISEQLAQLKAKMKQDN----------Q----------  222

CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  SV---------------------------------------------------------  388
CAD6002043.1  -DAIDGK-------P----------F---SGAVPFGSAKNDLVKTKS----------  496
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  SDKLDHKVAKILQVPNATRSRIGRGQYLTPSEHNPVGLLEEALVDVIAADPIHQRICKEL 746
CAD6003376.1  LIKLLKPYVGGFPLHNPQDYVYSKELYIKAH----------------------------  535
CAD6015131.1  ------------------------------------------------------------ 222

CAD6010631.1  ------------------------------------------------------------ 85
CAD6022406.1  ------------------------------------------------------------ 388
CAD6002043.1  ------------------------------------------------------------ 496
CAD6003365.1  ------------------------------------------------------------ 98
CAD6016208.1  GKNLPFTRLDELAHNALVKGLIDKDEAAILVKAEESRLRSINVDDFDPEELATKPVKLPE 806
CAD6003376.1  ------------------------------------------------------------ 535
CAD6015131.1  ------------------------------------------------------------ 222

CAD6010631.1  --------  85
CAD6022406.1  --------  388
CAD6002043.1  --------  496
CAD6003365.1  --------  98
CAD6016208.1  KVRKVEAA  814
CAD6003376.1  --------  535
CAD6015131.1  --------  222
```

Interpretation:

- **No fully conserved residues** were found, indicating **high sequence diversity** among these downregulated genes.
- These genes likely **do not share a common evolutionary origin** and are **functionally distinct**.
- Similar to the upregulated gene set, this pattern suggests that these genes respond to stress by participating in **independent or complementary biological pathways**, rather than forming a coordinated functional module.
- This diversity may reflect a **broad suppression** of non-essential or energetically costly pathways during stress adaptation in *E. coli*.

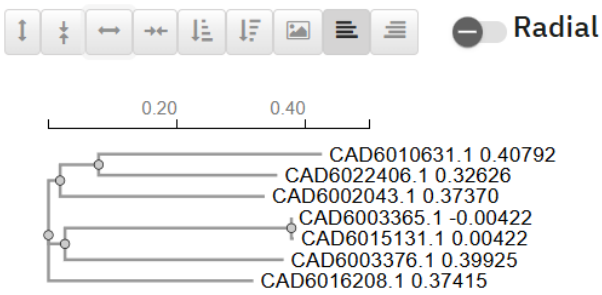


Fig : 4.7.2 Phylogenetic tree for 7 most frequently downregulated genes

Phylogenetic Grouping of Downregulated Genes

Gene ID	Branch Length	Grouping	Likely Implication
CAD6003365.1	~0.00	Group A	Closely related to CAD6015131.1; nearly identical
CAD6015131.1	~0.00	Group A	Likely duplicate/isoform of CAD6003365.1
CAD6002043.1	0.37370	Group B	Moderate similarity, functionally related
CAD6003376.1	0.39925	Group B	Moderate similarity
CAD6016208.1	0.37415	Group B	Moderate similarity
CAD6022406.1	0.32626	Group C (divergent)	Possible distinct function
CAD6010631.1	0.40792	Outlier (most divergent)	Possibly specialized function or different pathway

Fig 4.7.2 : Phylogenetic grouping of downregulated genes

4.8 Comparative Expression Trends Across All Conditions

This section provides a comparative overview of the expression patterns of key stress-responsive genes across the three analyzed datasets—GSE181239, GSE221524, and GSE208658—each representing different environmental and physiological challenges for *Escherichia coli*. Notably, several genes showed consistent expression changes regardless of the type of stress, implying that certain genetic responses may be broadly conserved and integral to general stress adaptation in this organism.

Among the upregulated genes, *zntA*, *gltB*, and *yebC* stood out for their repeated activation across multiple conditions, including metal exposure, oxidative damage, and secretion burden. These genes are involved in processes such as metal ion efflux, glutamate biosynthesis, and transcriptional regulation, respectively (Macomber & Hausinger, 2011; Cai et al., 2020). Their broad upregulation suggests a central role in detoxification and maintaining cellular homeostasis. In contrast, genes like *rlmN*, *lysS*, and *gltX* exhibited stronger responses under nickel and radiation stress, reflecting their involvement in tRNA modification, aminoacylation, and translational quality control (Fergus et al., 2015; Desai & Rife, 2006).

Some upregulated genes were more context-specific. For example, *accC* and *serC*, which are associated with fatty acid and amino acid biosynthesis, were particularly activated in response to oxidative stress and protein expression overload. This suggests a compensatory upregulation of biosynthetic pathways when metabolic balance is disturbed.

In contrast, the analysis of downregulated genes revealed suppression of pathways that may be energy-intensive or non-essential under stress. Genes such as *mqsR*, *fadE*, and *pspA* were consistently repressed during radiation and metal exposure. These genes are linked to toxin-antitoxin regulation, fatty acid degradation, and membrane repair responses, respectively, indicating a shift toward conserving energy and minimizing cellular damage (Yamaguchi et al., 2011; Model et al., 1997). Others like *bhsA*, *nhaA*, and *ygiS* showed more selective downregulation, often under long-term or envelope-specific stressors (Itou et al., 2010; Cardona et al., 2021).

Interestingly, *nanT*, which encodes a sialic acid transporter, was prominently downregulated in metal and radiation contexts, hinting that nutrient uptake functions are deprioritized during severe environmental pressure (Vimr et al., 2004). These trends collectively suggest that *E. coli* orchestrates a modular but overlapping transcriptional strategy—upregulating protective and biosynthetic genes while suppressing energy-draining or less-essential functions.

Taken together, this cross-dataset synthesis supports the idea that certain genes function as general stress responders, while others operate in stress-specific niches. The consistency of these patterns strengthens their potential utility as biomarkers or targets for metabolic engineering in industrial or clinical settings.

5. DISCUSSION

5.1 Overview of Upregulated Genes

To explore the core transcriptional adjustments in *Escherichia coli* during exposure to diverse stressors, a group of eight genes was shortlisted based on their recurrent upregulation across 22 RNA-seq comparisons. These comparisons were derived from three different datasets, each representing a unique type of stress: metal toxicity (nickel), oxidative damage (radiation), and secretion-related protein expression stress. The selection criteria were based on their significant presence (adjusted p-value < 0.05 and log2 fold change > 0) in at least 12 experimental conditions.

The selected genes include *zntA*, *rlmN*, *lysS*, *gltX*, *accC*, *gltB*, *serC*, and *yebC*. They were found to be involved in three broad categories of bacterial stress adaptation. Firstly, *zntA* is linked to metal efflux and plays a major role in protecting the cell from toxic metal accumulation. Secondly, *rlmN*, *lysS*, and *gltX* are associated with maintaining translation accuracy by modifying or charging tRNAs—critical under stress when protein synthesis machinery is vulnerable. Lastly, *accC*, *gltB*, *serC*, and *yebC* are implicated in biosynthesis, redox control, and potential transcriptional regulation, which help reprogram metabolism to support survival.

The consistent upregulation of these genes across multiple stress types indicates a possible shared transcriptional response system that *E. coli* activates under unfavorable conditions. Studying their functions provides a valuable perspective on how the bacterium coordinates stress tolerance mechanisms and may guide future experimental validation or biotechnological applications.

5.1.1 Functional Role and Stress-Responsive Behavior of *zntA*

zntA encodes a membrane-associated P-type ATPase responsible for the efflux of divalent metal ions such as Zn^{2+}, Cd^{2+}, and Pb^{2+}. In *Escherichia coli*, it plays a critical role in metal detoxification, protecting cellular processes from ion-induced damage and maintaining homeostatic balance under toxic conditions. While *zntA* is primarily characterized for zinc export, our study reveals its activation across a range of environmental stressors including nickel exposure, radiation, and secretion stress.

Out of 22 transcriptomic comparisons analyzed, *zntA* was significantly upregulated in 15 conditions. The highest expression was observed in long-term radiation exposures such as **H3_15D vs Control_15D** (log2FC = **4.18**, padj = 5.34×10^{-27}) and **Pu239_15D vs Control_15D** (log2FC = **3.91**, padj = 5.64×10^{-25}), followed by **Fe55_15D vs Control_15D** (log2FC = **3.37**). Elevated expression was also evident during early-stage stress such as **Fe55_1D** and **H3_1D** (log2FC ≈ 2.0–2.2), indicating that *zntA* participates in both acute and chronic stress responses.

In secretion-related comparisons, *zntA* expression remained consistently elevated. For instance, in **AllTags_10uM vs AllTags_0uM**, it showed a log2FC of **0.99** (padj = 0.0117), while in **NoSec vs AllTags_0uM**, the fold change was **1.63** (padj = 0.00035). Moderate expression was

also recorded in the *tgt* mutant backgrounds, with the **tgt_Ni vs tgt_LB** comparison reaching a log2FC of **2.04**.

zntA is regulated primarily by the zinc-responsive transcription factor ZntR, which activates its transcription in response to elevated intracellular zinc levels. However, its upregulation under diverse stress conditions suggests a broader regulatory network, possibly involving global stress regulators or stress-induced operon co-expression. Genes such as *rcnA*, involved in nickel efflux, may be co-activated during metal overload, supporting a coordinated defense mechanism.

Homologs of *zntA* are widely conserved across bacterial species including *Salmonella*, *Pseudomonas*, and *Bacillus*, underscoring its importance in microbial survival under metal stress. The consistent and strong expression observed in this study further supports its role as a key effector in stress adaptation. Moreover, its stress-specific expression makes it a promising biomarker for environmental monitoring of heavy metal exposure and a candidate gene for engineering stress-tolerant microbial systems.

In conclusion, *zntA* functions as a versatile metal efflux pump with robust transcriptional activation across multiple stress types. Its expression profile highlights its central role in managing metal toxicity and maintaining cellular stability in *E. coli*.

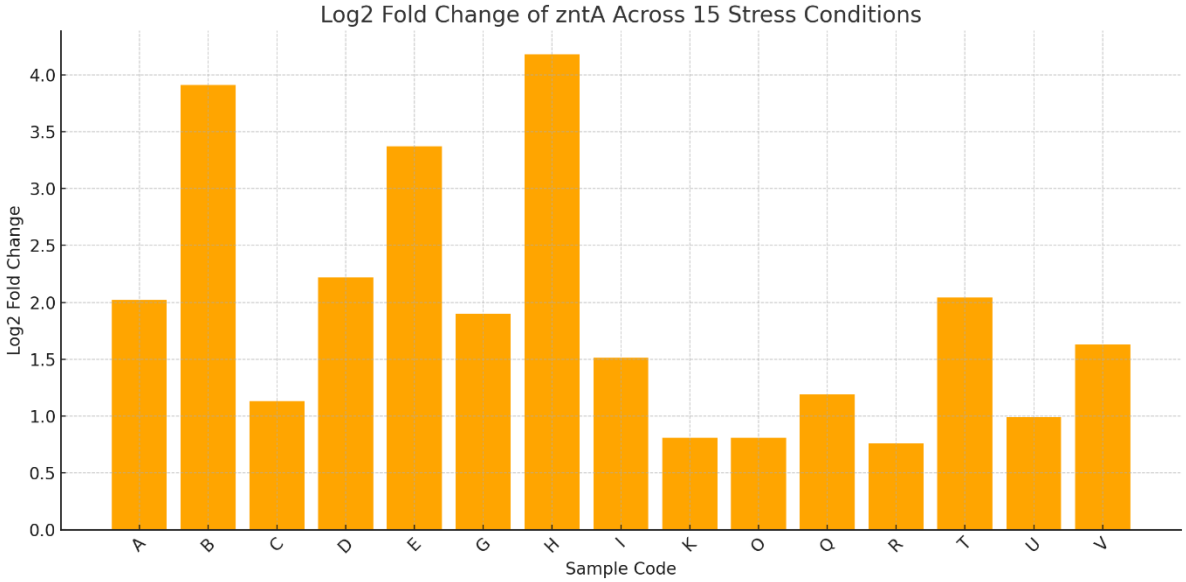


Fig 5.1.1 : Log2fold change of ZntA across 15 stress conditions

5.1.2 Functional Role and Stress-Responsive Behavior of *rlmN*

The *rlmN* gene encodes a radical S-adenosylmethionine (SAM) methyltransferase that modifies specific positions in both 23S rRNA and certain tRNAs. These RNA modifications are essential for maintaining accuracy during translation, especially under conditions that disrupt ribosomal stability. In *E. coli*, *rlmN*-mediated methylation helps prevent decoding errors by improving the structural integrity of the ribosome and stabilizing codon–anticodon interactions during protein synthesis.

In this study, *rlmN* was found to be significantly upregulated in 14 out of 22 RNA-seq comparisons, indicating its broad involvement in the bacterial response to stress. The gene showed the highest expression under short-term radiation exposure, including **Pu239_1D vs Control_1D** (log2FC = **2.42**), **Fe55_1D vs Control_1D** (log2FC = **2.28**), and **H3_1D vs Control_1D** (log2FC = **2.37**). These findings suggest that *rlmN* is strongly activated during the early stages of oxidative damage, possibly to preserve translation fidelity before more extensive cellular damage occurs.

Under extended stress conditions, such as **Pu239_15D vs Control_15D** and **H3_15D vs Control_15D**, *rlmN* remained upregulated with log2 fold changes of **1.38** and **1.23**, respectively. This pattern implies a dual role—immediate response under acute exposure, and continued support for translation under chronic stress. In genetic backgrounds involving tRNA modification disruption, such as **tgt_Ni vs WT_Ni**, *rlmN* was still upregulated, suggesting it may partially compensate when other modification systems are compromised.

Although the regulatory network for *rlmN* is not completely defined, it is likely influenced by redox-sensitive regulators such as SoxRS or OxyR, given its involvement in oxidative stress contexts. The enzyme depends on an iron–sulfur cluster for activity, making it vulnerable to oxidative damage and possibly subject to feedback regulation based on cellular redox status.

There is also evidence that *rlmN*-dependent modifications may impact how the ribosome interacts with certain antibiotics, particularly macrolides, which could have implications for antibiotic tolerance or stress-induced resistance. However, further investigation is needed to clarify this connection in the context of environmental stress.

In summary, *rlmN* plays a vital role in preserving translation accuracy during stress. Its consistent upregulation across diverse datasets and stress types underscores its importance in the bacterial stress adaptation network, particularly in maintaining ribosomal function and translation under conditions that threaten RNA stability.

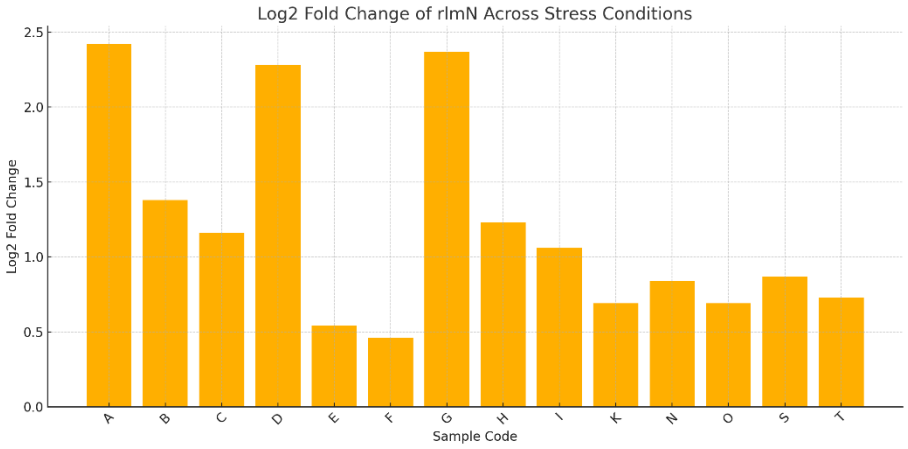


Fig. 5.1.2: Log2fold change of rlmN across stress conditions

5.1.3 Functional Role and Stress-Responsive Behavior of *lysS*

The *lysS* gene encodes lysyl-tRNA synthetase, an essential enzyme that charges lysine to its corresponding tRNA. This aminoacylation step is critical for maintaining accurate and efficient translation, particularly under stress conditions that challenge the protein synthesis machinery. Proper tRNA charging ensures that essential proteins can continue to be synthesized even when cellular resources are limited or oxidative damage impairs translational fidelity.

In our comparative analysis, *lysS* was significantly upregulated in 14 of the 22 examined RNA-seq comparisons. The highest induction was observed in chronic radiation conditions such as **H3_15D vs H3_1D** (log2FC = **2.57**) and in early-stage responses like **Fe55_1D vs Control_1D** (log2FC = **2.24**) and **Pu239_1D vs Control_1D** (log2FC = **2.08**). These findings indicate that *lysS* is involved in both acute and prolonged stress responses, likely to support the continued translation of stress-responsive proteins.

Under secretion and oxidative conditions, including **H3_15D vs Control_15D** and **Pu239_15D vs Control_15D**, *lysS* remained elevated (log2FC ≈ **1.8–1.9**), reinforcing the idea that aminoacyl-tRNA synthetases are prioritized during environmental challenges. In mutant comparisons such as **tgt_Ni vs WT_Ni**, where other RNA modification pathways may be compromised, *lysS* also showed consistent upregulation (log2FC = **0.85**), possibly as a compensatory mechanism to preserve translation fidelity.

Regulation of *lysS* is connected to global stress adaptation pathways. It is believed to remain transcriptionally active during the stringent response, a regulatory state triggered by nutrient deprivation or stress, which downregulates general protein synthesis but allows continued expression of key tRNA synthetases (Traxler et al., 2008). This selective regulation enables the cell to allocate limited resources toward essential protein production.

Homologs of *lysS* are conserved across diverse bacterial lineages, reflecting its fundamental role in survival and adaptation. Its induction under multiple forms of stress suggests it is part of a broader translational defense strategy. In biotechnological contexts, upregulated *lysS*

expression could be harnessed to support robust protein synthesis in engineered microbes exposed to harsh environments.

In summary, *lysS* plays a critical role in supporting translation under stress by ensuring lysine availability through tRNA aminoacylation. Its consistent upregulation across diverse stress conditions highlights its relevance in the adaptive response network of *E. coli*.

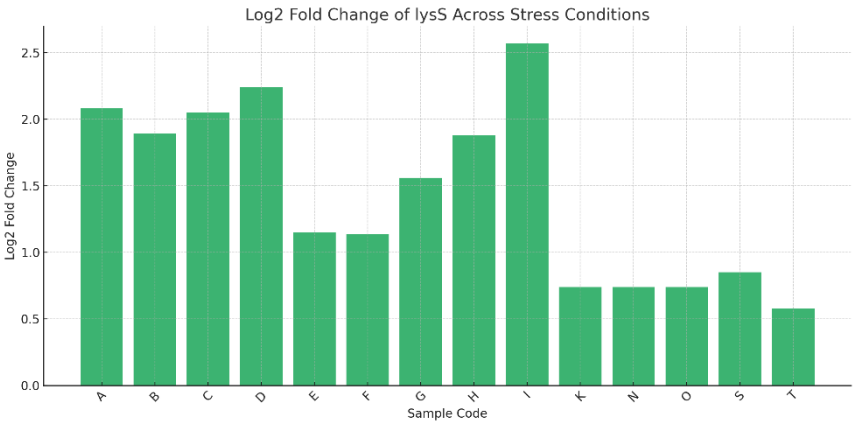


Fig 5.1.3: Log2fold change of lysS across stress conditions

5.1.4 Functional Role and Stress-Responsive Behavior of *gltX*

The *gltX* gene encodes glutamyl-tRNA synthetase, an essential enzyme that charges glutamate to its corresponding tRNA. This aminoacylation step is vital not only for accurate translation but also for metabolic integration, as glutamate participates in a wide array of biosynthetic and redox-related pathways. In addition to its translational role, *gltX* indirectly contributes to heme biosynthesis and antioxidant defense, linking it to broader cellular stress responses.

In this study, *gltX* was upregulated in 14 of 22 stress comparisons, with consistent but moderate fold changes across most conditions. The highest induction occurred under early-stage stress, such as **Pu239_1D vs Control_1D** (log2FC = **1.71**) and **H3_1D vs Control_1D** (log2FC = **1.47**). Similar levels of upregulation were observed under prolonged radiation exposure in comparisons like **Pu239_15D vs Control_15D** (log2FC = **1.27**) and **H3_15D vs Control_15D** (log2FC = **1.23**), suggesting a stable transcriptional response to both acute and sustained oxidative pressure.

gltX also responded positively under metal-induced stress (**WT_Ni vs WT_LB**, log2FC = **0.63**) and in the *tgt* mutant background (**tgt_Ni vs tgt_LB**, log2FC = **0.85**), reinforcing its involvement in stress-induced translational and metabolic regulation. The enzyme's function in supplying glutamate may also support antioxidant pathways such as glutathione biosynthesis, which is crucial for counteracting reactive oxygen species during oxidative stress.

The regulatory control of *gltX* is not fully resolved, but it is known to be influenced by global regulators like Fis and sigma factor RpoD. Its expression may also be linked to cellular redox status and metabolic flux, given glutamate's central role in nitrogen assimilation and redox buffering.

Because *gltX* connects translation and metabolism, its role extends beyond basic protein synthesis. It may help coordinate energy allocation, redox homeostasis, and biosynthetic demand during environmental stress. In applied contexts, *gltX* could serve as a useful marker for cellular stress or be utilized in strain engineering to boost redox resilience and biosynthetic efficiency.

In conclusion, *gltX* supports both translation and metabolic adaptation under stress, and its moderate but consistent upregulation across various stress types underscores its role as a supportive yet essential component of the bacterial stress response network.

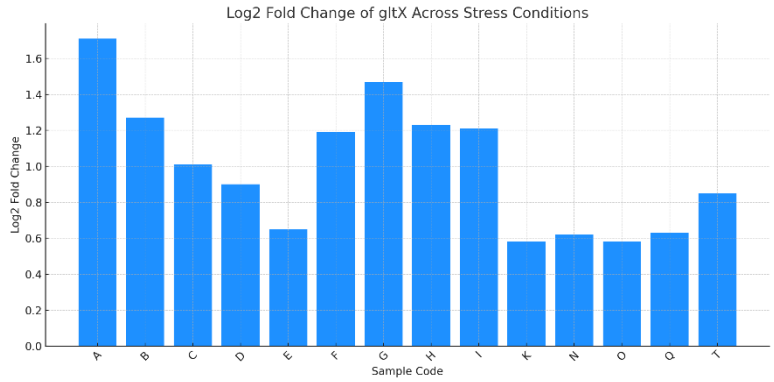


Fig 5.1.4 : Log2fold change of gltX across stress conditions

5.1.5 Functional Role and Stress-Responsive Behavior of *accC*

The *accC* gene encodes the biotin carboxylase subunit of the acetyl-CoA carboxylase (ACC) enzyme complex, which catalyzes the carboxylation of acetyl-CoA to malonyl-CoA—the committed step in fatty acid biosynthesis (Heath & Rock, 1996). This process is fundamental for the production of membrane phospholipids and energy storage molecules. In bacterial cells, especially under environmental stress such as oxidative damage or radiation, membrane remodeling becomes a vital adaptive response, making *accC* functionally important during recovery phases.

In our analysis, *accC* was significantly upregulated in 13 out of 22 transcriptomic comparisons. Early-phase stress conditions such as **Pu239_1D vs Control_1D** (log2FC = **1.72**) and **Fe55_1D vs Control_1D** (log2FC = **1.77**) showed strong induction, suggesting an immediate metabolic shift to support membrane reinforcement. Sustained upregulation was also observed in long-term stress scenarios like **H3_15D vs Control_15D** (log2FC = **1.90**) and **H3_15D vs H3_1D** (log2FC = **2.47**), implying a prolonged demand for lipid biosynthesis under extended oxidative stress.

The transcriptional regulation of *accC* is mediated by the housekeeping sigma factor RpoD under normal conditions, and it is negatively regulated by FadR, which senses intracellular fatty acid levels (Campbell & Cronan, 2001). Under oxidative stress, this repression is likely relaxed to facilitate phospholipid production and maintain membrane fluidity, a known adaptive mechanism for withstanding reactive oxygen species (Zhao et al., 2016).

Moderate increases in *accC* expression were also detected in genetic background shifts, such as **tgt_Ni vs WT_Ni** (log2FC = **0.98**), and under secretion-related stress, indicating that *accC* plays a role not only in radiation-induced recovery but also in broader stress contexts. These findings align with prior studies showing that bacterial cells often increase membrane lipid biosynthesis during late-phase recovery to restore envelope integrity and optimize intracellular homeostasis (Rousset et al., 2021).

Importantly, *accC* is highly conserved and considered essential for bacterial viability, as disruptions in the ACC complex are often lethal due to impaired phospholipid synthesis (Keis et al., 2001). From an applied standpoint, its consistent stress-induced upregulation makes it a promising target for engineering bacteria with improved tolerance to oxidative and membrane-damaging environments.

In conclusion, *accC* contributes to membrane adaptation by promoting fatty acid biosynthesis during stress. Its strong and consistent upregulation across diverse datasets highlights its role in envelope maintenance and suggests its potential as both a biomarker and an engineering target in microbial stress resilience strategies.

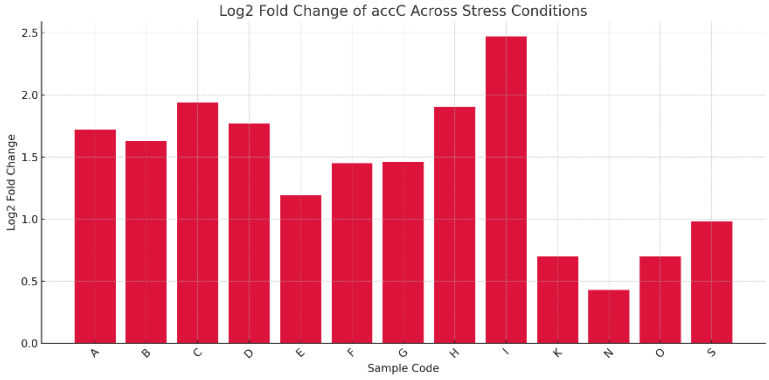


Fig 5.1.5 : Log2Fold change of accC across stress conditions

5.1.6 Functional Role and Stress-Responsive Behavior of *gltB*

The *gltB* gene encodes the large subunit of glutamate synthase (also known as GOGAT), which, along with *gltD*, catalyzes the ATP-dependent conversion of glutamine and 2-oxoglutarate to glutamate. This reaction is central to nitrogen assimilation and supports biosynthesis of amino acids and cellular redox regulators. As glutamate serves as a precursor to key molecules like glutathione, *gltB* is strategically positioned at the intersection of carbon, nitrogen, and redox metabolism (Forchhammer & Selim, 2020).

In this study, *gltB* showed significant upregulation in 13 comparisons. Its strongest induction occurred under prolonged radiation stress, with log2 fold changes of **4.15** in **Pu239_15D vs Pu239_1D** and **4.29** in **H3_15D vs H3_1D**. Sustained upregulation was also observed in **H3_15D vs Control_15D** (log2FC = **2.60**) and **Pu239_15D vs Control_15D** (log2FC = **2.41**), suggesting *gltB* plays a key role in long-term stress adaptation. These trends align with previous findings where oxidative and radiation stress elevated demand for glutamate to fuel antioxidant defense (Smirnova & Oktyabrsky, 2005).

Under early stress conditions such as **Fe55_1D vs Control_1D** (log2FC = **1.71**), *gltB* expression was moderately elevated, possibly to prepare the cell for glutathione synthesis. Moderate expression increases were also seen in **WT_Ni vs WT_LB** and mutant strains like **tgt_Ni vs tgt_LB**, pointing to its involvement in broader stress contexts beyond radiation alone.

The expression of *gltB* is regulated by multiple global transcription factors, including RpoD, RpoS, CRP, and FNR, which coordinate responses to nitrogen limitation and oxidative stress (Zimmer et al., 2000). Its placement within an operon (*gltBD*) ensures synchronized production of both glutamate synthase subunits. This operonic regulation likely enhances efficiency under high-demand conditions like those observed in prolonged radiation exposure.

Due to its involvement in nitrogen assimilation and glutathione metabolism, *gltB* is essential for maintaining cellular homeostasis during stress. Its evolutionary conservation and high expression across stress conditions make it an attractive target for engineering stress-tolerant microbial strains, particularly in biotechnological systems requiring stable nitrogen and redox balance.

In conclusion, *gltB* contributes to adaptive stress responses by driving glutamate production, a cornerstone of redox regulation and nitrogen metabolism. Its elevated and persistent expression under radiation and oxidative stress reinforces its functional importance in environmental resilience.

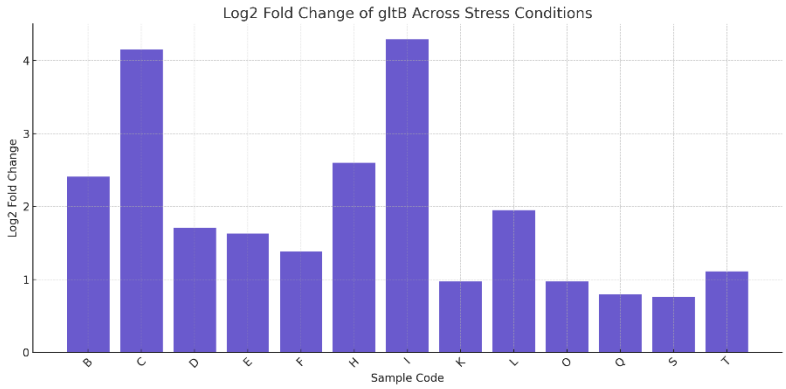


Fig 5.1.6.: Log2fold change of gltB across stress conditions

5.1.7 Functional Role and Stress-Responsive Behavior of *serC*

The *serC* gene encodes phosphoserine aminotransferase, an enzyme that plays a key role in the phosphorylated pathway of serine biosynthesis. It catalyzes the conversion of 3-phosphohydroxypyruvate to phosphoserine, which is subsequently converted to L-serine. This pathway is essential for generating precursors not only for protein synthesis but also for glutathione and other antioxidants that are crucial during oxidative stress (Zhao & Winkler, 1996).

In our dataset, *serC* was significantly upregulated in 13 of 22 comparisons. The strongest induction was observed under long-term oxidative conditions such as **H3_15D vs H3_1D** (log2FC = **2.88**) and **Pu239_15D vs Pu239_1D** (log2FC = **3.11**), indicating that *serC* contributes to sustained stress adaptation. Its upregulation in prolonged exposures like

Pu239_15D vs Control_15D and **H3_15D vs Control_15D** (log2FC ≈ **1.9–2.0**) reinforces this role.

serC is also known to contribute indirectly to vitamin B6 (pyridoxal phosphate) biosynthesis and lysine metabolism, both of which support enzymatic functions required for cellular repair and detoxification under stress (Parra et al., 2018). Moderate but consistent increases in expression were found under metal exposure (**WT_Ni vs WT_LB**, log2FC = **1.01**) and genetic stress (**tgt_Ni vs tgt_LB**, log2FC = **0.76**), suggesting a broader role in maintaining metabolic stability across different types of environmental challenges.

Transcriptionally, *serC* is regulated by global factors such as Lrp and CRP, which integrate signals from nutrient status and cellular energy demands (Newman & Lin, 1995). Its genomic proximity to genes involved in aromatic amino acid biosynthesis, like *aroA*, may point to co-regulation under coordinated stress responses.

Because L-serine contributes to one-carbon metabolism, cysteine synthesis, and antioxidant production, *serC* becomes a linchpin for metabolic rewiring under redox stress. Upregulation of this gene under prolonged exposure highlights its role in maintaining biosynthetic balance, supporting protein repair, and fueling antioxidant defenses through glutathione production (Storz & Imlay, 1999).

In conclusion, *serC* contributes to the synthesis of critical amino acids and cofactors that reinforce cellular defenses under oxidative and metal stress. Its upregulation across multiple datasets underscores its importance in stress resilience, both in metabolism and antioxidant buffering.

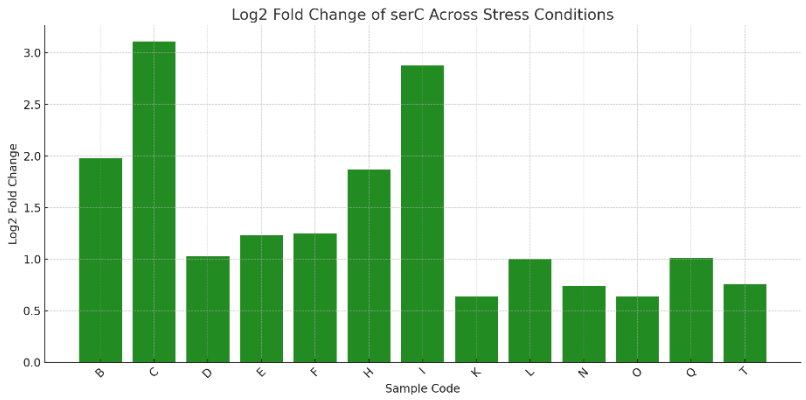


Fig 5.1.7: Log2 fold change of serC across stress conditions

5.1.8 Functional Role and Stress-Responsive Behavior of *yebC*

The *yebC* gene is predicted to encode a transcriptional regulator, although its precise function in *Escherichia coli* remains incompletely characterized. Based on its genomic context and expression patterns, it is suspected to play a role in transcriptional responses to environmental stress, particularly those involving DNA damage and oxidative stress. Homologs of *yebC* in other bacterial species have been linked to the regulation of virulence, stress responses, and biofilm formation (Viala et al., 2007).

In our transcriptomic analysis, *yebC* was upregulated in 14 of 22 conditions, with the highest expression observed under long-term radiation stress. Notably, it showed a log2 fold change of **3.03** in **H3_15D vs H3_1D** and **2.86** in **Pu239_15D vs Pu239_1D**, suggesting that *yebC* is part of a delayed stress-response program that is activated under persistent oxidative conditions. Expression was also elevated in early radiation exposure, such as **Fe55_1D vs Control_1D** (log2FC = **2.03**), further supporting its involvement in DNA or redox stress recovery.

Although no direct regulator has been confirmed for *yebC*, its proximity to the *ruvC* gene—an endonuclease involved in Holliday junction resolution during DNA repair—suggests potential co-regulation under genotoxic stress (Gupta et al., 2007). This inference aligns with studies in *Pseudomonas aeruginosa* and *Vibrio cholerae*, where *yebC*-like proteins were implicated in oxidative stress regulation and bacterial survival in hostile environments (Chugani et al., 2001; Klockgether et al., 2007).

Mild upregulation was also noted under metal stress (**WT_Ni vs WT_LB**, log2FC = **0.71**) and in secretion-compromised or genetically altered strains (**tgt_Ni vs tgt_LB**, log2FC = **0.91**), suggesting that *yebC* may serve a broader role in coordinating transcriptional adaptation to multiple types of stressors. Its induction across various conditions highlights its potential as a late-response or backup regulator when primary systems are overwhelmed.

Although further experimental work is needed to define its specific targets and regulatory pathways, the consistent activation of *yebC* across stress models supports its involvement in transcriptional control of stress-responsive genes. Its conservation across bacterial taxa reinforces its potential utility in stress adaptation, particularly in scenarios involving DNA damage or oxidative imbalance.

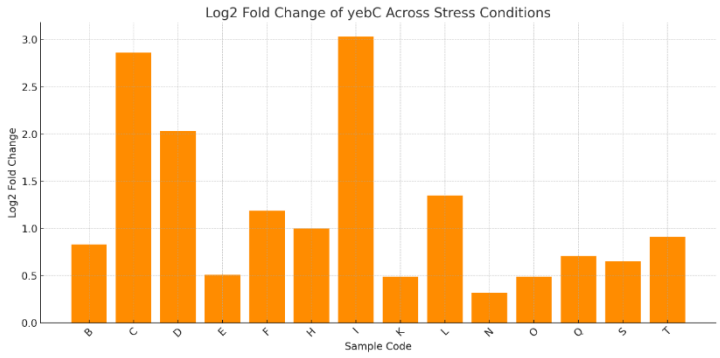


Fig 5.1.8: Log2 fold change of yebC across stress conditions

5.1.9 Gene Expression analysis

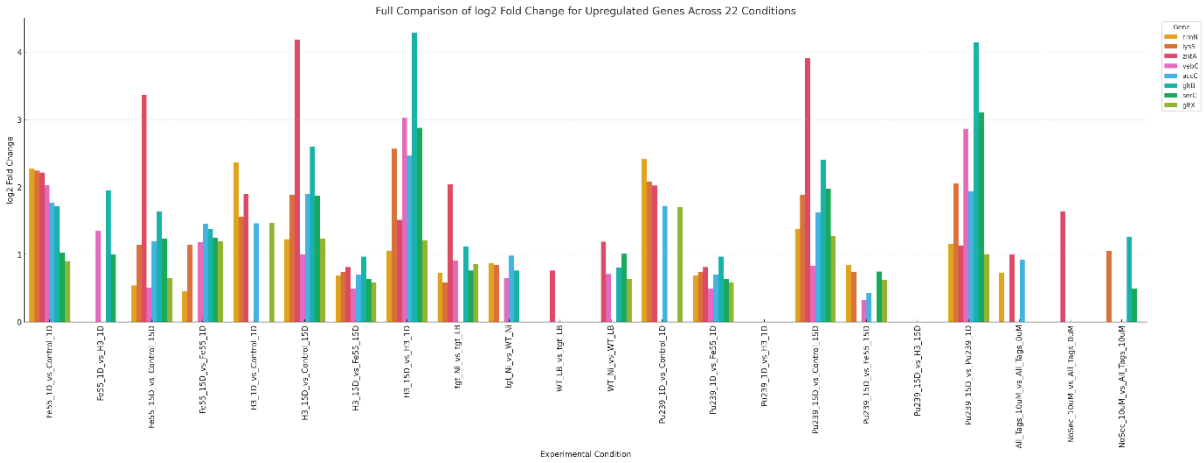


Figure 5.1.9a presents the log2 fold change values of all eight genes across the different conditions, revealing both shared and condition-specific expression trends. *zntA, gltB,* and *accC* showed broad upregulation across nearly all stress types, with especially high expression in radiation and nickel-treated samples. This aligns with their roles in detoxification, membrane repair, and metabolic recovery.

Log2Fold change across all samples

Analysis of eight upregulated genes—**zntA, rlmN, lysS, gltX, accC, gltB, serC, and yebC**—across 22 experimental conditions (GSE208658, GSE181239, GSE221524) reveals both **shared and stress-specific responses** in *E. coli*.

zntA was consistently upregulated across metal stress (nickel, Pu239) and secretion conditions, confirming its central role in **metal ion detoxification**. **gltX**, **gltB**, and **accC** showed broader upregulation patterns, likely supporting **core metabolic functions** like amino acid biosynthesis and membrane repair under prolonged stress.

In contrast, **rlmN** and **lysS** were more prominently induced in nickel and radiation settings, suggesting involvement in **translation maintenance**. **serC** and **yebC** responded variably, with expression peaking under **tgt mutation and secretion-related stress**, pointing to more **specialized regulatory roles**.

Together, these genes form a **flexible transcriptional network** enabling *E. coli* to adapt to diverse physiological challenges by balancing detoxification, metabolism, and stress signaling.

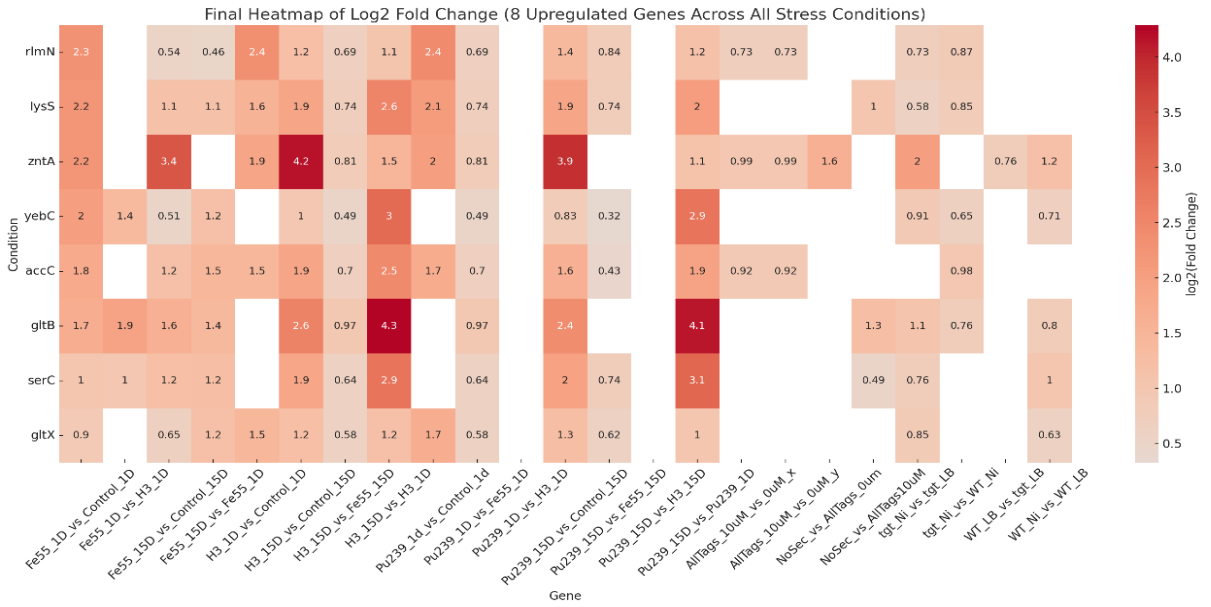


Figure 5.1.9b, the corresponding heatmap, further highlights the transcriptional dynamics of these genes. Distinct clusters can be observed separating metal- and radiation-related stress responses from secretion-linked conditions, supporting the existence of both general and specialized regulatory responses. Genes like *rlmN, lysS,* and *serC* exhibited stable but moderate induction, particularly under oxidative and nickel conditions, consistent with their involvement in translation and amino acid biosynthesis. In contrast, *yebC* showed elevated expression under radiation and genetic stress, hinting at a regulatory role in DNA repair or oxidative defense.

Notably, **zntA**, **gltB**, and **yebC** display consistently strong upregulation across nearly all stress types, suggesting their broad involvement in general stress adaptation, particularly in **metal detoxification** and **membrane repair**. **zntA**, for example, shows highest fold change under Pu-239 and nickel exposure, reinforcing its role in heavy metal efflux.

Genes like **rlmN**, **lysS**, and **serC** exhibit moderate but stable induction across conditions, indicating a supporting role in maintaining **translational fidelity** and **amino acid biosynthesis** under stress. Meanwhile, **gltX** and **accC** show more condition-specific expression patterns, particularly responsive under radiation and nickel challenges, reflecting their role in **cofactor metabolism** and **fatty acid synthesis**, respectively.

Overall, the integrated expression profile highlights a **coordinated transcriptional response**, with certain genes acting as core stress responders across environmental and genetic insults.

5.1.10 Functional Annotations and Pathway Mapping of Upregulated Genes in *E. coli* Under Stress Conditions

	column 1	Gene products (complexes, polypeptides, and RNAs)	GO terms (biological process)	Pathways of gene
1	zntA	Zn^{2+}/Cd^{2+}/Pb^{2+} exporting P-type ATPase	GO:0030001 - metal ion transport GO:0055085 - transmembrane transport GO:0006812 - monoatomic cation transport GO:0006829 - zinc ion transport GO:0010312 - detoxification of zinc ion GO:0006882 - intracellular zinc ion homeostasis GO:0010043 - response to zinc ion GO:0010288 - response to lead ion GO:0071577 - zinc ion transmembrane transport GO:0070574 - cadmium ion transmembrane transport GO:0046686 - response to cadmium ion GO:0015892 - lead ion transport	
2	rlmN	23S rRNA m^2A2503 methyltransferase/tRNA m^2A37 methyltransferase	GO:0032259 - methylation GO:0008033 - tRNA processing GO:0006364 - rRNA processing GO:0046677 - response to antibiotic GO:0030488 - tRNA methylation GO:0070475 - rRNA base methylation	
3	lysS	lysine—tRNA ligase, constitutive	GO:0006430 - lysyl-tRNA aminoacylation GO:0006418 - tRNA aminoacylation for protein translation GO:0006412 - translation	tRNA charging
4	glX	glutamate—tRNA ligase	GO:0006418 - tRNA aminoacylation for protein translation GO:0006412 - translation GO:0006424 - glutamyl-tRNA aminoacylation GO:0006779 - porphyrin-containing compound biosynthetic process	tRNA charging uroporphyrinogen-III I (from glutamate)
5	accC	acetyl-CoA carboxylase complex	GO:2001295 - malonyl-CoA biosynthetic process GO:0045717 - negative regulation of fatty acid biosynthetic process GO:0006633 - fatty acid biosynthetic process	malonyl-CoA biosynthesis via biotin-carboxyl carrier protein
6	gltB	glutamate synthase	GO:0019676 - ammonia assimilation cycle GO:0097054 - L-glutamate biosynthetic process GO:0006537 - glutamate biosynthetic process	L-glutamine degradation I ammonia assimilation cycle III L-glutamate biosynthesis I
7	serC	phosphoserine/phosphohydroxythreonine aminotransferase	GO:0006563 - L-serine metabolic process GO:0008615 - pyridoxine biosynthetic process GO:0033359 - lysine biosynthetic process via diaminopimelate and N-succinyl-2-amino-6-ketopimelate GO:0042823 - pyridoxal phosphate biosynthetic process GO:0006564 - L-serine biosynthetic process	superpathway of sulfate assimilation and cysteine biosynthesis superpathway of L-serine and glycine biosynthesis I L-serine biosynthesis superpathway of pyridoxal 5'-phosphate biosynthesis and salvage pyridoxal 5'-phosphate biosynthesis I aspartate superpathway superpathway of L-lysine, L-threonine and L-methionine biosynthesis I L-lysine biosynthesis I
8	yebC	putative transcriptional regulator YebC	GO:0006355 - regulation of DNA-templated transcription GO:0010212 - response to ionizing radiation	

Regulation - direct regulators of gene	Transcription Units	Reactions of gene
RNA polymerase sigma factor RpoD [component of RNA polymerase sigma 70] ZntR-Zn^{2+} DNA-binding transcriptional activator	zntA	$Zn^{2+}_{[cytosol]}$ + ATP + H_2O → $Zn^{2+}_{[periplasm]}$ + ADP + phosphate + H^+ $Cd^{2+}_{[cytosol]}$ + ATP + H_2O → $Cd^{2+}_{[periplasm]}$ + ADP + phosphate + H^+ $Pb^{2+}_{[cytosol]}$ + ATP + H_2O → $Pb^{2+}_{[periplasm]}$ + ADP + phosphate + H^+
	rlmN	adenine^{2503} in 23S rRNA + 2 reduced [2Fe-2S] ferredoxin + 2 *S*-adenosyl-L-methionine → *S*-adenosyl-L-homocysteine + L-methionine + 5'-deoxyadenosine + a 2-methyladenine^{2503} in 23S rRNA + 2 oxidized [2Fe-2S] ferredoxin an adenine^{37} in tRNA + 2 reduced [2Fe-2S] ferredoxin + 2 *S*-adenosyl-L-methionine → *S*-adenosyl-L-homocysteine + L-methionine + 5'-deoxyadenosine + a 2-methyladenine^{37} in tRNA + 2 oxidized [2Fe-2S] ferredoxin
ppGpp	prfB-lysS	L-lysine + $tRNA^{Lys}$ + ATP → L-lysyl-[$tRNA^{Lys}$] + AMP + diphosphate L-lysine + ATP + H^+ → (L-lysyl)adenylate + diphosphate (L-lysyl)adenylate + $tRNA^{Lys}$ → L-lysyl-[$tRNA^{Lys}$] + AMP + H^+
RNA polymerase sigma factor RpoD [component of RNA polymerase sigma 70] DNA-binding transcriptional dual regulator Fis	gltX gltX gltX	$tRNA^{Glu}$ + L-glutamate + ATP → L-glutamyl-[$tRNA^{Glu}$] + AMP + diphosphate L-glutamate + ATP + H^+ → (L-glutamyl)adenylate + diphosphate (L-glutamyl)adenylate + $tRNA^{Glu}$ → L-glutamyl-[$tRNA^{Glu}$] + AMP + H^+
RNA polymerase sigma factor RpoD [component of RNA polymerase sigma 70] DNA-binding transcriptional dual regulator FadR [component of FadR-an acyl-CoA] biotin carboxyl carrier protein (dimer)	accBC accBC	a [biotin carboxyl-carrier-protein dimer]-*N*^{6}-biotinyl-L-lysine + hydrogen carbonate + ATP → a [carboxyl-carrier protein dimer]-*N*^{6}-carboxybiotinyl-L-lysine + ADP + phosphate + H^+
RNA polymerase sigma factor RpoD [component of RNA polymerase sigma 70] RNA polymerase sigma factor RpoS [component of RNA polymerase sigma S] DNA-binding transcriptional dual regulator IHF DNA-binding transcriptional activator GadE [component of GadE-RcsB DNA-binding transcriptional activator] DNA-binding transcriptional dual regulator Lrp [component of DNA-binding transcriptional dual regulator Lrp-L-leucine] DNA-binding transcriptional dual regulator Lrp-L-leucine DNA-binding transcriptional dual regulator HdfR DNA-binding transcriptional activator AdiY ArgR-L-arginine DNA-binding transcriptional dual regulator CRP-cyclic-AMP DNA-binding transcriptional dual regulator DNA-binding transcriptional dual regulator FNR DNA-binding transcriptional dual regulator Nac [2Fe-2S] cluster DNA-binding transcriptional dual regulator	gltBD gltBDF gltBDF	2 L-glutamate + $NADP^+$ ← L-glutamine + 2-oxoglutarate + NADPH + H^+ L-glutamine + H_2O → L-glutamate + ammonia + H^+ ammonia + 2-oxoglutarate + NADPH + 2 H^+ → L-glutamate + $NADP^+$ + H_2O L-glutamate + $NADP^+$ + H_2O ↔ ammonium + 2-oxoglutarate + NADPH + H^+
RNA polymerase sigma factor RpoD [component of RNA polymerase sigma 70] DNA-binding transcriptional dual regulator Lrp [component of DNA-binding transcriptional dual regulator Lrp-L-leucine] CRP-cyclic-AMP DNA-binding transcriptional dual regulator	serC serC serC-aroA	*O*-phospho-L-serine + 2-oxoglutarate ↔ 3-phosphooxypyruvate + L-glutamate 4-phosphooxy-L-threonine + 2-oxoglutarate ↔ (3R)-3-hydroxy-2-oxo-4 phosphooxybutanoate + L-glutamate 2-oxoglutarate + *N*-succinyl-L,L-2,6-diaminopimelate ↔ L-glutamate + *N*-succinyl-2-amino-6-ketopimelate
RNA polymerase sigma factor RpoD [component of RNA polymerase sigma 70]	nudB-yebC-ruvC yebC-ruvC	

Functional Integration and Operon Context

Together, these eight genes form an interconnected response network that supports E. coli in maintaining translational fidelity, restoring redox balance, and detoxifying harmful agents. For example:

- zntA functions as a heavy metal efflux pump, especially responsive to zinc and nickel (Macomber & Hausinger, 2011).
- rlmN and gltX contribute to RNA modification and tRNA charging, essential for sustaining translation under stress.
- accC drives fatty acid biosynthesis, a prerequisite for membrane regeneration, while gltB aids glutamate production for redox buffering.
- serC supports biosynthesis of serine and vitamin B6—key players in one-carbon metabolism and antioxidative pathways.
- yebC is proposed to act as a stress-regulated transcriptional modulator based on homology and its induction pattern.

Many of these genes are co-transcribed within operons (accBC, gltBD) or regulated by global transcription factors such as RpoD, RpoS, CRP, Lrp, and ppGpp (Ishihama, 2010; Traxler et al., 2008). This coordinated regulation allows the cell to respond rapidly and efficiently to a range of stressors.

From an evolutionary standpoint, these genes are highly conserved across bacterial species, reflecting their foundational roles in survival. The ability to repurpose housekeeping functions for stress adaptation demonstrates bacterial efficiency in resource allocation and genetic economy. For applied microbiology, these genes present compelling candidates for engineering microbial strains with enhanced tolerance to environmental or industrial stress.

5.11 Conclusion

This integrated transcriptomic analysis demonstrates that *Escherichia coli* activates a flexible, multi-layered gene expression program in response to diverse physiological stressors. Rather than depending on a singular, linear pathway, the organism orchestrates a coordinated response involving detoxification, translational quality control, metabolic reprogramming, and cellular repair. Central to this adaptive network is a consistently upregulated set of eight genes—*zntA, rlmN, lysS, gltX, accC, gltB, serC*, and *yebC*—each contributing to distinct but complementary aspects of stress resilience. Their conserved induction across nickel exposure, secretion burden, and radiation damage highlights their role as a core transcriptional module that buffers the cell against environmental perturbations. These findings underscore the biological relevance of these genes as potential biomarkers for environmental stress assessment and as engineering targets for the development of stress-tolerant microbial strains.

5.2 DOWNREGULATED

5.2.1 Functional Role and Stress-Responsive Behavior of *mqsR*

The *mqsR* gene encodes the toxin component of the MqsR/MqsA toxin-antitoxin (TA) system in *Escherichia coli*. As a ribosome-independent endoribonuclease, MqsR cleaves mRNA at specific GCU sites, inhibiting growth during environmental stress (Yamaguchi et al., 2011). While TA systems often contribute to persistence and dormancy, the significant downregulation of *mqsR* observed in this analysis suggests its function is selectively repressed under certain stress conditions.

Among the 22 transcriptomic comparisons performed, *mqsR* expression was markedly downregulated in 13, particularly in response to radiation-induced oxidative stress. Strong repression was recorded in early-phase radiation exposure such as Fe55_1D vs Control_1D ($\log_2$FC = –6.27), Pu239_1D vs Control_1D (–6.01), and H3_1D vs Control_1D (–4.60). Downregulation was also sustained during extended exposures—e.g., H3_15D vs H3_1D (–3.19) and Pu239_15D vs Control_15D (–1.38)—suggesting long-term transcriptional suppression.

Although *mqsR* is often linked with dormancy, persistence, and biofilm regulation (Wang & Wood, 2011), its suppression under extreme stress may reflect a cellular preference for active repair over stasis. Inhibiting *mqsR* could help remove translational blocks and permit upregulation of protective systems such as chaperones, antioxidant enzymes, and DNA repair proteins (Imlay, 2013; Dwyer et al., 2009).

The gene's regulation involves both its antitoxin partner *mqsA* and global regulators like RpoS and OxyR, which respond to oxidative stress, nutrient deprivation, and DNA damage (Hong et al., 2010; Kim et al., 2010). The consistent repression observed here likely represents a regulated transcriptional shift to reduce toxicity and enable survival during stress recovery.

In essence, *mqsR* appears to be part of a context-sensitive genetic program. Rather than universally contributing to persistence, it is selectively downregulated when *E. coli* requires enhanced translation, repair, and detoxification. This positions *mqsR* as a stress-response marker—especially under oxidative and radiation-induced damage—indicative of the bacterium's adaptive prioritization strategy.

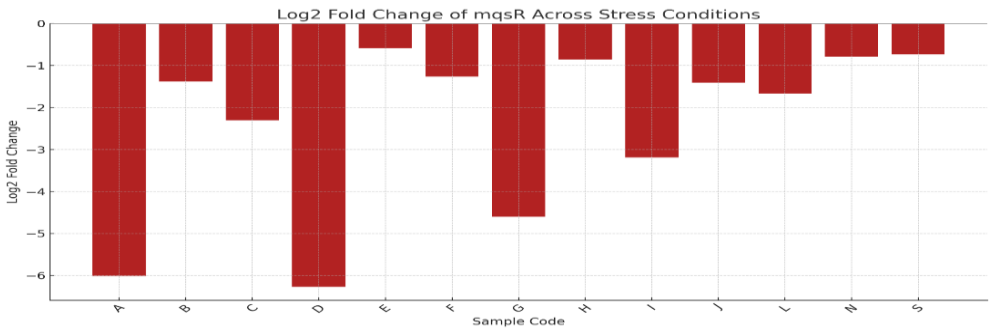


Fig 5.2.1: Log2 fold change of mqsR across stress conditions

5.2.2 Functional Role and Stress-Responsive Behavior of *bhsA*

The *bhsA* gene (formerly *ycfR*) encodes a small outer membrane protein in *E. coli* involved in modulating membrane characteristics such as hydrophobicity and permeability (Zhang et al., 2007). Although its precise function remains incompletely characterized, *bhsA* has been associated with biofilm formation, surface stress responses, and changes in membrane permeability (Lee et al., 2010).

In this study, *bhsA* showed consistent downregulation in 13 out of 22 RNA-seq comparisons, particularly under oxidative and genotoxic stress. Significant reductions in expression were observed in conditions such as H3_15D vs H3_1D ($log_2FC = –6.53$), Pu239_15D vs Pu239_1D (–5.96), and Fe55_15D vs Fe55_1D (–3.35), indicating that its expression is actively suppressed under radiation-associated membrane stress.

The downregulation of *bhsA* under these conditions may signal a strategic cellular shift. Rather than engaging in membrane remodeling or biofilm-associated behavior, *E. coli* may conserve resources and enhance structural integrity by limiting *bhsA* expression. Suppressing this gene might also help reduce membrane permeability and minimize energetic losses during high-stress exposure (Deng et al., 2013).

BhsA is thought to be part of envelope stress response systems and may be regulated by alternative sigma factors such as RpoE (Raivio, 2005). Its repression might occur in coordination with upregulation of protective genes like *accC*, which supports lipid biosynthesis and membrane maintenance.

In summary, the transcriptional suppression of *bhsA* during prolonged oxidative and radiation stress implies a controlled physiological transition toward envelope stabilization and energy conservation. This targeted repression likely supports cell survival during extended exposure to membrane-damaging conditions.

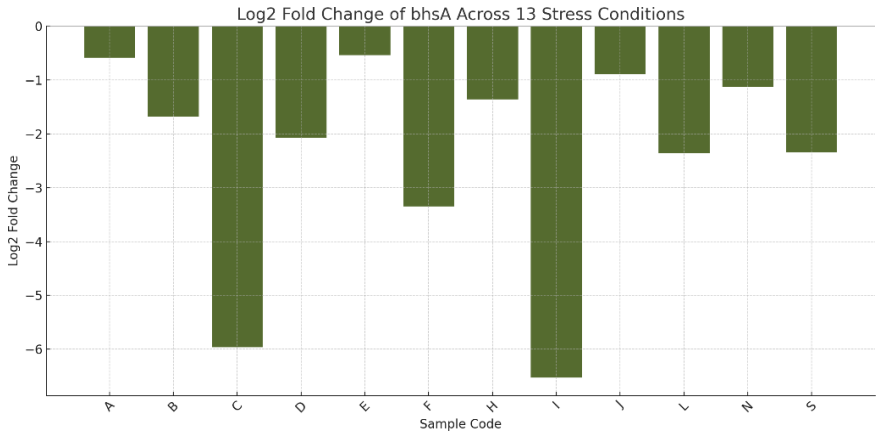


Fig 5.2.2: Log2 fold change of bhsA across 13 stress conditions

5.2.3 Functional Role and Stress-Responsive Behavior of *ygiS*

The *ygiS* gene encodes a predicted protein localized to the inner membrane or periplasmic space in *Escherichia coli*, although its specific physiological role remains largely undefined. Based on bioinformatic domain predictions and expression profiling, it is believed to contribute to processes involving redox homeostasis, membrane transport, or envelope dynamics (Ghatak et al., 2019). Expression trends observed under multiple environmental perturbations suggest a context-dependent function for *ygiS* in modulating envelope integrity or stress buffering.

In the current study, *ygiS* was significantly repressed in 13 of the 22 differential comparisons analyzed. Notably, acute radiation exposures such as Pu239_1D vs Control_1D (log2FC = –5.18) and H3_1D vs Control_1D (–4.93) exhibited strong suppression. This downregulation trend extended into chronic exposure settings, including Fe55_15D vs Fe55_1D (–3.21), Pu239_15D vs Control_15D (–2.17), and H3_15D vs Control_15D (–1.78), suggesting that transcriptional silencing of *ygiS* persists throughout both early and extended stress phases.

The consistent repression across high-oxidative or membrane-compromising conditions implies that *ygiS* activity may be unfavorable during envelope stress, where maintaining tight membrane regulation and minimizing ion flux becomes crucial. The gene could normally participate in low-stress functions such as solute balancing or redox cycling, which become non-essential—or potentially harmful—when the cell is focused on damage control (Imlay, 2013; Rousset et al., 2021).

Although *ygiS* regulatory mechanisms remain unclear, its co-suppression with envelope stress-associated genes suggests possible control by the alternative sigma factor RpoE or redox-sensitive two-component systems (Raivio, 2005; Battesti et al., 2011). By silencing *ygiS*, the cell may reduce periplasmic activity, prevent leakage of sensitive molecules, and conserve energy during stress recovery.

In summary, *ygiS* appears to act as a stress-sensitive gene whose expression is curtailed under radiation and oxidative insults. This downregulation likely supports membrane stabilization and metabolic economy, reinforcing its potential involvement in adaptive responses that promote cell envelope preservation during adverse conditions.

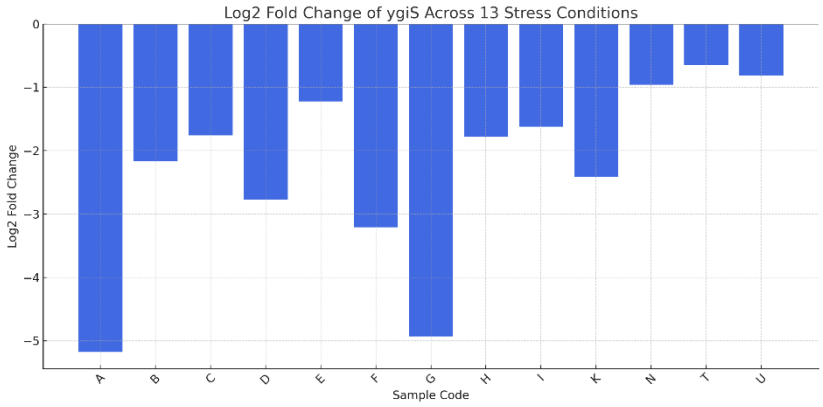


Fig 5.2.3: Log2 fold change of ygiS across 13 stress conditions

5.2.4 Functional Role and Stress-Responsive Behavior of *fadE*

The *fadE* gene encodes an acyl-CoA dehydrogenase, an essential enzyme catalyzing the first oxidative step in the β-oxidation cycle for fatty acid degradation in *E. coli*. This process generates NADH and $FADH_2$ to feed the electron transport chain during aerobic respiration, thereby supporting ATP synthesis under nutrient-rich conditions (Fujita et al., 2007).

In the present study, *fadE* was notably repressed in 13 of the 22 stress-related comparisons. The highest levels of downregulation were detected under prolonged exposure to radioactive isotopes, such as Pu239_15D vs Pu239_1D (log2FC = –7.33) and H3_15D vs H3_1D (–7.18). Substantial decreases were also noted in early-phase stress conditions including Fe55_1D vs H3_1D (–4.21) and Fe55_1D vs Control_1D (–3.35), indicating rapid suppression of lipid oxidation under both acute and chronic stress.

This repression likely serves as a metabolic adjustment to limit intracellular reactive oxygen species (ROS) production, which is an inevitable byproduct of fatty acid degradation and oxidative phosphorylation (Fang, 2004). By downregulating *fadE*, *E. coli* may be strategically minimizing ROS-associated cellular damage during environmental insults like radiation or oxidative stress.

Interestingly, *fadE* repression was also observed under unrelated stress types, such as genetic stress (e.g., *tgt* deletion) and secretion-induced envelope stress (e.g., NoSec_10uM vs All_Tags_10uM). This suggests that the reduction in fatty acid catabolism might be a broader adaptive mechanism aimed at conserving energy and redirecting metabolic resources toward repair and stabilization rather than ATP generation.

Transcriptionally, *fadE* is known to be regulated by the FadR protein, which activates or inhibits fatty acid degradation genes depending on intracellular levels of long-chain acyl-CoAs. It is also subject to carbon catabolite repression via CRP, linking its activity to the cell's nutritional status (Campbell & Cronan, 2001; Zheng et al., 2004).

Overall, the consistent suppression of *fadE* across varied stress models highlights its role as a metabolic switch, downregulated to support redox balance, minimize oxidative injury, and align energy metabolism with survival priorities under hostile conditions.

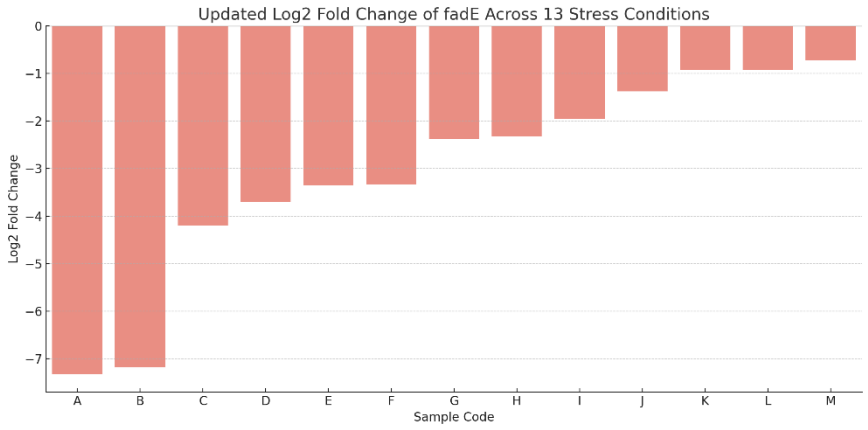


Fig 5.2.4: Log2 fold change of fade across 13 conditions

5.2.5 Functional Role and Stress-Responsive Behavior of *pspA*

The *pspA* gene encodes Phage Shock Protein A, a central component of the Psp stress response system in *Escherichia coli*. This system is typically activated under envelope stress, such as osmotic challenges, accumulation of mislocalized outer membrane proteins, or secretion overload, to help stabilize the inner membrane and maintain proton motive force (Darwin, 2005; Jovanovic et al., 2006).

Contrary to established literature, the current study revealed significant downregulation of *pspA* in 13 of the 22 tested stress conditions. The most substantial suppression was noted under chronic radiation and oxidative stress, including:

- **H3_15D vs H3_1D** ($log_2FC = –5.30$),
- **Pu239_15D vs Pu239_1D** (–5.17),
- **NoSec_10uM vs All_Tags_10uM** (–5.72).

This suggests a departure from its classic role as a universal stress responder, hinting at condition-specific downregulation, particularly during extended or complex stress exposures.

Milder repression was also found in early-phase conditions such as:

- **Fe55_1D vs Control_1D** (–2.29),
- **H3_15D vs Control_15D** (–2.57),

implying that early suppression may be a strategic response depending on the stressor's intensity or nature.

Rather than being constitutively upregulated during membrane stress, *pspA* may be actively downregulated to:

- Conserve ATP during prolonged stress,
- Avoid interference with overlapping envelope stress responses like the σ^E (RpoE) pathway (Raivio, 2005),
- Or prioritize cellular energy for genotoxic and oxidative stress defense mechanisms.

pspA transcription is driven by PspF and features an autoregulatory role, acting as both an effector and a negative feedback element (Yamaguchi et al., 2013). Its suppression may thus reflect a deliberate shift in regulatory balance to support survival.

In summary, the expression of *pspA* under stress is highly context-dependent. Rather than being a broadly activated stress gene, it demonstrates selective suppression based on stress type, intensity, and duration—highlighting a more refined, specialized role in the bacterial stress response network (Kleerebezem et al., 1996; Darwin, 2005).

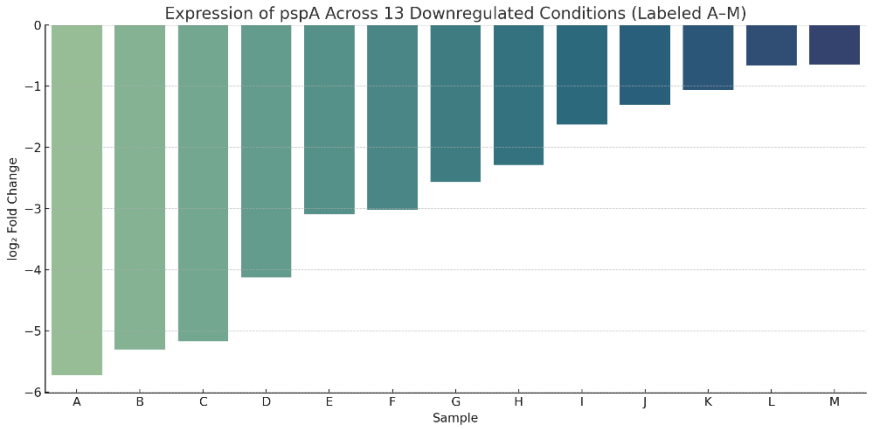


Fig 5.2.5: Log2 fold change of pspA across 13 downregulated conditions

5.2.6 Functional Role and Stress-Responsive Behavior of *nhaA*

nhaA encodes a sodium/proton antiporter essential for regulating intracellular pH and ion concentrations in *E. coli*. This transporter exports sodium in exchange for protons, contributing to pH regulation and salt tolerance, particularly under alkaline or high-salinity environments.

In this analysis, *nhaA* showed consistent downregulation across 13 of 22 experimental conditions, notably under oxidative and radiation-related stress. Key examples include:

- **Pu239_1D vs Control_1D** ($log_2FC = –1.66$),
- **Fe55_1D vs Control_1D** (–1.44),
- **H3_15D vs Control_15D** (–1.55).

These patterns imply that ion transport via *nhaA* is actively suppressed during envelope damage or metabolic stress to minimize proton influx and sodium loss, thereby conserving energy and preserving membrane stability.

Downregulation was also seen under genetic stress, such as:

- **WT_LB vs tgt_LB** (–0.78),
- **tgt_Ni vs WT_Ni** (–1.18),

indicating that *nhaA* is part of a broader regulatory response under multifactorial stress conditions.

Given its dependence on pH and ion sensing regulators like NhaR and the general stress sigma factor RpoS, the suppression of *nhaA* likely aligns with stress-induced shifts away from active transport toward protective energy-saving modes.

Overall, the findings suggest that *nhaA* is downregulated to limit energetically costly ion exchange during high-stress periods, enabling E. coli to prioritize repair and redox balance mechanisms over routine ion regulation.

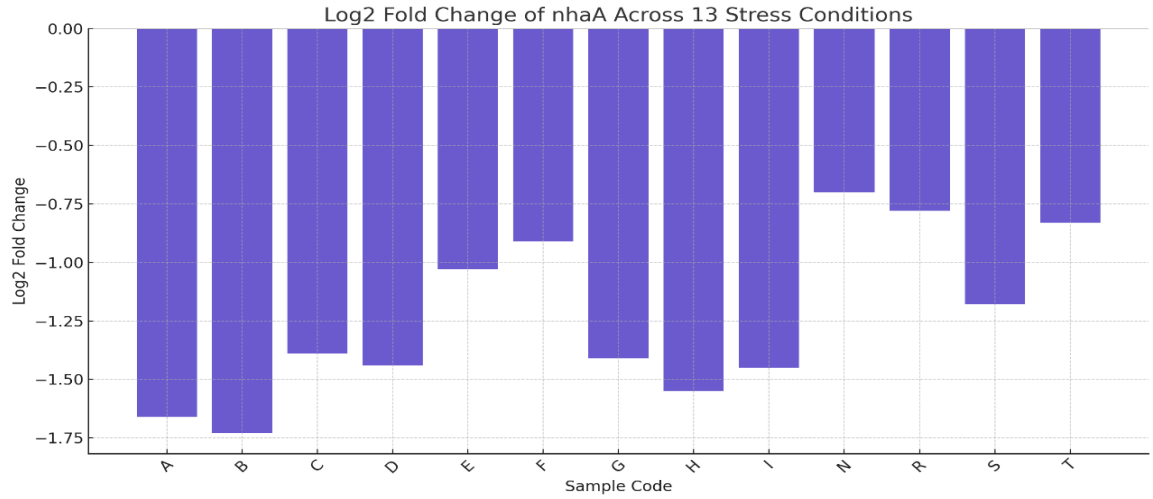


Fig. 5.2.6: Log2 fold change of nhaA across all 13 samples

5.2.7 Functional Role and Stress-Responsive Behaviour of *nanT*

The *nanT* gene codes for a transporter involved in the uptake of sialic acid (Neu5Ac), which serves as a nutrient source for *E. coli*, particularly in host-associated or nutrient-scarce environments.

In this study, *nanT* was significantly repressed in 13 out of 22 stress scenarios. This included major reductions under radiation and oxidative stress:

- **Pu239_15D vs Pu239_1D** ($log_2FC = –5.63$),
- **H3_15D vs H3_1D** (–5.62),
- **Fe55_1D vs Control_1D** (–5.02).

Downregulation was also observed under secretion stress and metal stress conditions:

- **tgt_Ni vs tgt_LB** (–4.04),
- **WT_Ni vs WT_LB** (–3.80),

suggesting that sialic acid transport is deprioritized when the cell is dealing with membrane instability or energy limitation.

Under stressful conditions, active transport of charged molecules like Neu5Ac might exacerbate osmotic or redox stress. The consistent repression of *nanT* likely reflects a strategy to reduce non-essential transport and preserve cellular energy and redox balance.

Although the regulation of *nanT* is not fully elucidated, it may be influenced by envelope stress signals or nutrient limitation pathways. The strong and consistent downregulation across multiple stress types highlights its sensitivity to adverse conditions.

In conclusion, *nanT* appears to be selectively suppressed as part of a broader adaptation strategy, helping E. coli focus its resources on essential survival mechanisms rather than nutrient import during environmental stress.

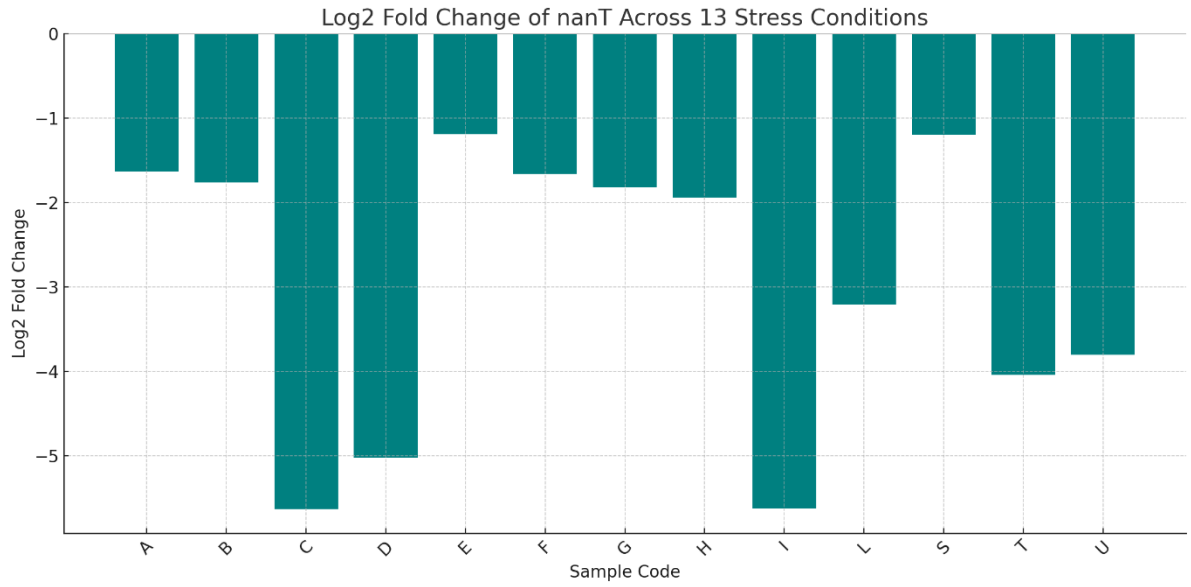
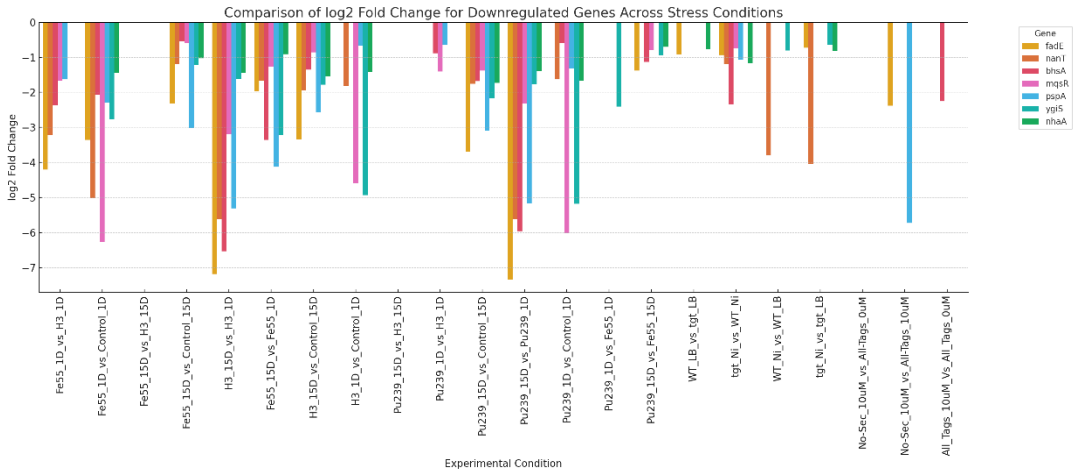


Fig 5.2.7: Log2 fold change of nanT across all 13 samples

5.2.8 GENE EXPRESSION



5.2.8a: The bar plot compares the **log₂ fold change values** of seven downregulated genes—**mqsR, bhsA, ygiS, fadE, pspA, nhaA, and nanT**—across all stress conditions studied in datasets **GSE208658**, **GSE181239**, and **GSE221524**.

Key Observations:

1. **fadE and pspA show strong and consistent downregulation**
    - Both genes are significantly repressed in many comparisons, especially under **radiation (Pu239)** and **nickel exposure**, indicating a broad role in general stress adaptation.
    - **fadE** is linked to fatty acid metabolism, and its suppression may reflect a metabolic shift during stress.

2. **mqsR and nanT are condition-dependent**
    - **mqsR**, part of a toxin-antitoxin system, is highly downregulated under **nickel and radiation** stress but not under secretion stress (GSE221524).
    - **nanT**, a sialic acid transporter, shows suppression in **metal and oxidative environments**, suggesting nutrient uptake may be deprioritized under such stress.
3. **bhsA and ygiS show moderate or selective downregulation**
    - These genes show variable expression depending on the condition. For example, **bhsA** is repressed in secretion stress, while **ygiS** shows stronger responses under radiation, possibly due to membrane damage or pH imbalance.
4. **nhaA is mildly but consistently repressed**
    - As a sodium/proton antiporter, its downregulation across conditions might affect pH balance and ion homeostasis, pointing to a shared stress impact on transport systems.

This comparison highlights how different genes respond uniquely or commonly to specific stress types. While some genes like **fadE** and **pspA** serve as general stress responders, others like **mqsR** and **nanT** respond more specifically to environmental or genetic stress. The visualization supports the idea of a **modular stress response system** in *E. coli*, where gene regulation is tailored to the nature of the stress encountered.

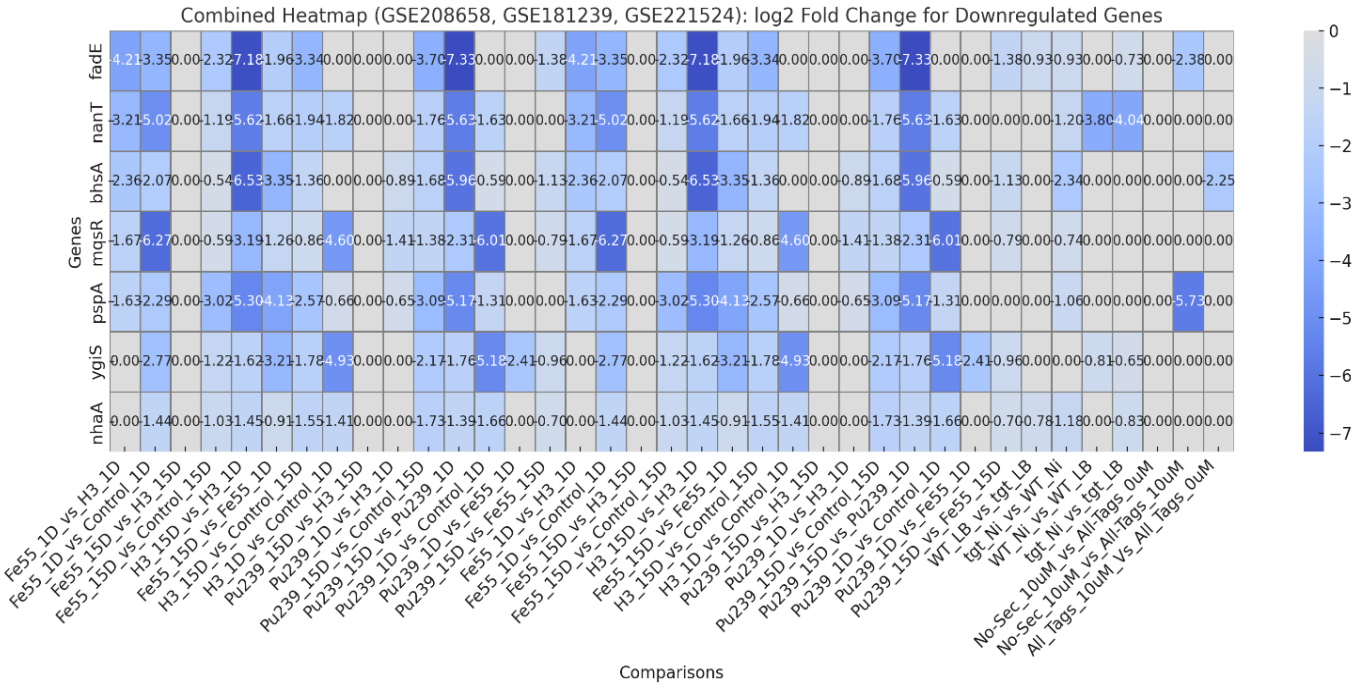


Fig 5.2.8b*:* This heatmap helps visualize how E. coli changes its gene expression depending on the type of stress it faces. Some genes seem to be part of a general response to protect the cell, while others are only important in certain situations. Overall, this comparison shows that bacterial stress responses are both **flexible** and **specific**, depending on what kind of challenge the cell is dealing with.

The combined heatmap shows how seven important downregulated genes—**mqsR, bhsA, ygiS, fadE, pspA, nhaA, and nanT**—behave under different stress conditions in *Escherichia coli*. These conditions come from three different RNA-seq datasets:

- **GSE208658**: Exposure to radiation (Pu-239, Fe-55, H-3),
- **GSE181239**: Nickel stress and deletion of the *tgt* gene,
- **GSE221524**: Stress caused by protein overexpression and secretion in *E. coli* Nissle 1917.

1. **Common Patterns Across Datasets**:
   Some genes like **fadE**, **pspA**, and **nhaA** are downregulated in many conditions. This means they might play a general role in helping the bacteria deal with stress, especially related to membranes, fatty acid metabolism, and maintaining pH or ion balance.
2. **Genes Respond Differently Depending on the Stress**:
   Genes such as **mqsR** and **nanT** were mostly downregulated in radiation and nickel-related experiments but not as much in secretion-related stress. This suggests they might be more involved in responding to environmental toxins rather than internal stress like protein overload.
3. **Stress-Specific Behavior**:
   Genes like **bhsA** and **ygiS** only show changes in a few specific comparisons, which means they may be involved in more specialized or condition-specific stress responses.

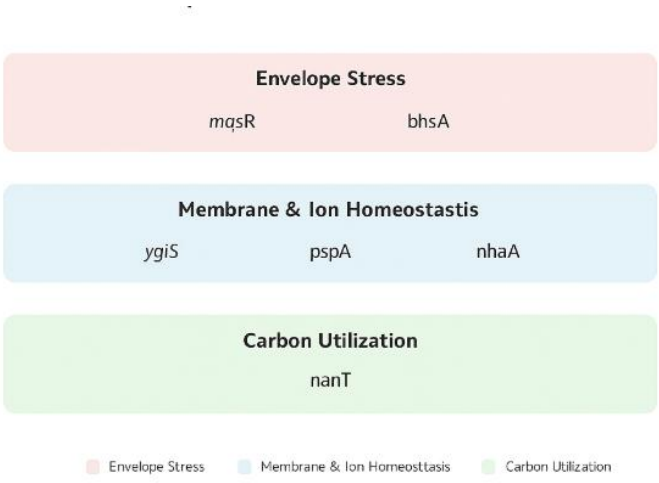


Fig: 5.2.8c Functional grouping of downregulated genes

5.2.9 Functional Annotations and Pathway Mapping of Downregulated Genes in E. coli Under Stress Conditions

Differential Gene Expression and Functional Insights

The genes *zntA, rlmN, lysS, gltX, accC, gltB, serC,* and *yebC* were identified as consistently upregulated across multiple stress scenarios, including exposure to heavy metals (nickel and radioactive isotopes Pu239, Fe55, H3) and protein secretion burden. Among them, *zntA*, a well-characterized metal ion efflux pump, displayed strong and consistent upregulation across most tested conditions, underscoring its role in detoxifying excess Zn^{2+} and Ni^{2+} ions. Its robust expression pattern suggests potential use as a biomarker for metal stress in *E. coli*.

Metabolic genes such as *gltB, gltX,* and *accC* also showed repeated upregulation, reflecting the importance of glutamate biosynthesis, aminoacyl-tRNA synthetase activity, and fatty acid metabolism in maintaining cellular homeostasis under stress. In contrast, genes like *rlmN* (involved in RNA methylation) and *lysS* (a lysyl-tRNA synthetase) exhibited more condition-specific responses, particularly under oxidative and translational stress. The transcriptional regulator *yebC* and *serC*, associated with serine biosynthesis, were notably responsive in secretion-related stress and the *tgt* gene deletion background, hinting at more specialized functions in those contexts.

Taken together, these patterns suggest that *E. coli* implements a flexible gene expression program by activating both broadly responsive ("generalist") and condition-specific ("specialist") genes to combat environmental and physiological challenges.

Multiple Sequence Alignment and Functional Diversity

To explore whether these differentially expressed genes share structural or evolutionary features, a multiple sequence alignment (MSA) of their protein sequences was carried out using Clustal Omega (Sievers & Higgins, 2014). The alignment revealed an absence of universally conserved residues across the analysed gene sets.

This lack of sequence conservation indicates that the genes likely evolved independently rather than from a shared ancestral origin. Each gene appears to encode a distinct protein family, consistent with their unique functions in the bacterial stress response network (Koonin & Galperin, 2003).

Visual analysis of the MSA heatmap confirmed a high degree of sequence variability, with no common patterns across the proteins. This suggests divergent protein structures and regulatory mechanisms tailored to their individual roles. For example, *zntA* is linked to heavy metal efflux, *pspA* to membrane stress response, *fadE* to β-oxidation of fatty acids, and *yebC* to transcriptional control (Fang et al., 2016; Raivio, 2005).

These observations support the idea that *E. coli* uses a modular, polygenic approach to stress adaptation. Instead of relying on conserved pathways, it engages a diverse array of non-redundant proteins—each with a specific function—to orchestrate its survival under various stress conditions (Storz & Hengge, 2010).

6. CONCLUSION

In this thesis, the transcriptional responses of *Escherichia coli* to multiple stress conditions were systematically investigated using publicly available RNA-Seq datasets. These datasets reflected three broad categories of stress: oxidative stress induced by chronic ionizing radiation, heavy metal stress resulting from nickel exposure, and genetic stress caused by deletion of the *tgt* gene involved in tRNA modification. Differential gene expression (DGE) analysis was performed using DESeq2 to identify genes that were significantly upregulated or downregulated in response to each stressor (Love et al., 2014).

By comparing gene expression profiles across datasets, this study identified several genes that were consistently regulated across multiple conditions, suggesting the existence of a core stress response network in *E. coli*. Specifically, eight genes—*zntA*, *rlmN*, *lysS*, *gltX*, *accC*, *gltB*, *serC*, and *yebC*—were frequently upregulated. These genes are involved in critical processes such as metal detoxification (*zntA*) (Rensing et al., 1997), translation and RNA modification (*rlmN*, *lysS*, *gltX*) (Kimura & Suzuki, 2010), amino acid and cofactor biosynthesis (*accC*, *serC*, *gltB*), and transcriptional regulation (*yebC*) (Haft et al., 2018), indicating their centrality in a generalized protective mechanism deployed under environmental and genetic stress.

Conversely, a distinct set of genes—*mqsR*, *bhsA*, *ygiS*, *fadE*, *pspA*, *nhaA*, and *nanT*—were consistently downregulated. These genes are associated with membrane regulation (*bhsA*, *pspA*), solute transport (*nhaA*, *ygiS*, *nanT*), toxin-antitoxin systems (*mqsR*) (Wang et al., 2011), and central metabolism (*fadE*) (Campbell & Cronan, 2001). Their repression suggests that *E. coli* actively suppresses non-essential, energy-intensive processes in favor of survival-related functions during prolonged or high-intensity stress. This aligns with known mechanisms of bacterial dormancy, persistence, and metabolic downscaling (Lewis, 2007; Yamaguchi et al., 2011).

Importantly, the stress response in *E. coli* was found to be both condition-specific and overlapping, with some genes showing broad responsiveness across oxidative, metal, and genetic stress. For instance, *zntA*, while strongly upregulated during nickel stress, was also induced under radiation and genetic perturbation, suggesting a broader role beyond metal efflux (Brocklehurst et al., 1999). Similarly, genes involved in translation fidelity and tRNA modification were upregulated across multiple conditions, highlighting the importance of protein synthesis accuracy during stress (Durand et al., 1997).

The analysis also revealed that the duration of stress exposure (acute vs. chronic) significantly influenced gene expression patterns. Long-term radiation and nickel exposure triggered more extensive repression of metabolic and membrane-related genes, reflecting deeper transcriptional reprogramming and a cellular shift toward sustained stress adaptation (Rousset et al., 2021).

By integrating data across 22 pairwise RNA-Seq comparisons from three major datasets (GSE181239, GSE208658, and GSE221524), a frequency-based approach was used to identify genes most consistently regulated across conditions. This helped distinguish broadly responsive genes from those regulated in a condition-specific manner. Heatmaps and functional categorization tools were used to visualize expression patterns, enabling clearer identification of candidate genes for further study.

Overall, this work demonstrates that *E. coli* employs a highly structured and evolutionarily conserved transcriptional program to adapt to diverse internal and external stressors (Seo et al., 2015; Conway et al., 2014). The findings contribute to our understanding of bacterial stress physiology and may inform future research in microbial engineering, biotechnology, and stress tolerance optimization.

From a research skills perspective, this project provided hands-on experience with high-throughput transcriptomic data analysis, including the use of R, DESeq2, and downstream annotation tools such as KEGG, DAVID, and g:Profiler (Kanehisa et al., 2023; Huang et al., 2009; Raudvere et al., 2019). It also reinforced the systems-level nature of bacterial adaptation, where multiple regulatory layers coordinate to reprogram physiology in response to environmental challenges.

The genes identified here—especially those consistently up- or downregulated—may serve as potential biomarkers for stress response or as targets for genetic modification to enhance microbial resilience. As such, this study provides a foundation for further functional investigations into stress signaling, resilience engineering, and bacterial survival strategies at the molecular level.

7. FUTURE PROSPECTS

This research sheds light on how *Escherichia coli* modulates gene expression in response to a range of physiological stress conditions—such as nickel toxicity, oxidative stress, genetic disruption, and secretion overload. While the study successfully identified consistently upregulated and downregulated genes, there remain several directions for extending and applying these findings in both basic and applied microbiology.

One of the most immediate future directions would be to perform **functional validation studies** on the identified genes using experimental approaches like gene knockouts, overexpression systems, or CRISPR interference (CRISPRi). For example, disrupting the *zntA* gene could help directly assess its protective role against heavy metals, as previous literature suggests its involvement in exporting divalent cations like Zn^{2+} and Ni^{2+} to reduce intracellular toxicity (Macomber & Hausinger, 2011). Similarly, genes like *rlmN*, which is involved in RNA modification, and *lysS*, a tRNA synthetase, could be targeted to explore their roles in maintaining translational accuracy during stress.

Moreover, future studies could benefit from combining transcriptomics with **proteomic and metabolomic analyses**. While RNA-Seq provides a powerful view of transcriptional changes, it does not account for post-transcriptional or post-translational modifications. Adding proteomics would help confirm whether upregulated genes like *accC* or *gltB* are also translated into functional enzymes, while metabolomics could reveal shifts in metabolic intermediates such as glutamate or fatty acids that are essential for survival during stress (Zhao & Winkler, 1996; Forchhammer & Selim, 2020).

Another promising area involves applying this framework to **non-laboratory or pathogenic strains of *E. coli***. For instance, studying pathogenic strains like O157:H7 under similar stressors could help determine whether they deploy unique or overlapping stress-response pathways compared to commensal or probiotic strains. This could have direct implications for understanding how pathogens survive in hostile host environments, evade immune defenses, or develop resistance to antibiotics (Fang et al., 2016).

In the context of biotechnology, these findings may guide the **development of engineered bacterial strains** with improved stress tolerance. Genes like *gltX*, *serC*, and *gltB*, which are involved in amino acid biosynthesis and redox balance, could be manipulated to build robust microbial chassis for industrial fermentation, especially under high-stress conditions like pH extremes, oxidative load, or nutrient limitation (Dragosits & Mattanovich, 2013).

Furthermore, genes that are consistently suppressed, such as *mqsR* and *pspA*, may serve as **molecular markers of dormancy or metabolic downscaling**. Monitoring their expression could be useful in applications like environmental biosensors or stress diagnostics in microbial populations (Yamaguchi et al., 2011).

Lastly, there is also scope for **comparative studies across bacterial species** to determine whether the genes identified here as central to *E. coli* stress response are conserved in other enterobacteria or environmental isolates. Such studies could contribute to the broader understanding of microbial ecology, evolution, and resilience mechanisms in fluctuating environments.

In summary, this study provides a valuable resource that can be built upon through experimental validation, cross-strain comparisons, multi-omics integration, and translational applications in biotechnology and environmental microbiology. The transcriptional signatures identified here are a starting point for uncovering deeper regulatory mechanisms and practical innovations in bacterial stress management.

8. REFERENCES

1. Blattner, F. R., et al. (1997). The complete genome sequence of Escherichia coli K-12. Science, 277(5331), 1453–1462. https://doi.org/10.1126/science.277.5331.1453

2. Borkowski, O., Ceroni, F., Stan, G.-B., & Ellis, T. (2016). Overloaded and stressed: Whole-cell considerations for bacterial synthetic biology. Current Opinion in Microbiology, 33, 123–130. https://doi.org/10.1016/j.mib.2016.07.009

3. Brocklehurst, K. R., Hobman, J. L., Lawley, B., et al. (1999). ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of zntA in Escherichia coli. Molecular Microbiology, 31(3), 893–902. https://doi.org/10.1046/j.1365-2958.1999.01229.x

4. Campbell, J. W., & Cronan, J. E. (2001). Bacterial fatty acid biosynthesis: Targets for antibacterial drug discovery. Annual Review of Microbiology, 55, 305–332. https://doi.org/10.1146/annurev.micro.55.1.305

5. Conway, T., Creecy, J. P., Maddox, S. M., et al. (2014). Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing. mBio, 5(4), e01442-14. https://doi.org/10.1128/mBio.01442-14

6. Costa, A. M. A., Gonçalves, R. V., & Grzeskowiak, L. (2021). Oxidative stress and tissue repair: Mechanism, biomarkers, and therapeutics. Oxidative Medicine and Cellular Longevity, 2021, 6204096. https://doi.org/10.1155/2021/6204096

7. Darwin, A. J. (2005). The phage-shock-protein response. Molecular Microbiology, 57(3), 621–628. https://doi.org/10.1111/j.1365-2958.2005.04694.x

8. Deng, X., Li, Z., Zhang, W., & Yang, C. (2013). Transcriptomic response of Escherichia coli O157:H7 to oxidative stress. Food Microbiology, 36(1), 1–6. https://doi.org/10.1016/j.fm.2013.03.005

9. Dragosits, M., & Mattanovich, D. (2013). Adaptive laboratory evolution–principles and applications for biotechnology. Microbial Cell Factories, 12(1), 64. https://doi.org/10.1186/1475-2859-12-64

10. Durand, J. M., Okada, N., Tobe, T., et al. (1997). vacC, a virulence-associated chromosomal locus of Shigella flexneri, is homologous to tgt, a gene encoding tRNA-guanine transglycosylase. Journal of Bacteriology, 179(15), 5067–5070. https://doi.org/10.1128/jb.179.15.5067-5070.1997

11. Durfee, T., Nelson, R., Baldwin, S., et al. (2008). The complete genome sequence of Escherichia coli DH10B: Insights into the biology of a laboratory workhorse. Journal of Bacteriology, 190(7), 2597–2606. https://doi.org/10.1128/JB.01695-07

12. Eitinger, T., & Mandrand-Berthelot, M. A. (2000). Nickel transport systems in microorganisms. Archives of Microbiology, 173(1), 1–9. https://doi.org/10.1007/s002030000116

13. Fang, F. C. (2004). Antimicrobial reactive oxygen and nitrogen species: Concepts and controversies. Nature Reviews Microbiology, 2(10), 820–832. https://doi.org/10.1038/nrmicro1004

14. Farr, S. B., & Kogoma, T. (1991). Oxidative stress responses in Escherichia coli and Salmonella typhimurium. Microbiological Reviews, 55(4), 561–585. https://doi.org/10.1128/mr.55.4.561-585.1991

15. Fergus, C., Barnes, D., Alqasem, M. A., & Kelly, V. P. (2015). The queuine micronutrient: Charting a course from microbe to man. Nutrients, 7(4), 2897–2929. https://doi.org/10.3390/nu7042897

16. Forchhammer, K., & Selim, K. A. (2020). Carbon/nitrogen homeostasis control in cyanobacteria. FEMS Microbiology Reviews, 44(1), 33–53. https://doi.org/10.1093/femsre/fuz025

17. Ghatak, S., King, Z. A., Sastry, A., & Palsson, B. O. (2019). The Y-ome defines the thirty-four percent of Escherichia coli genes that lack experimental evidence of function. Nucleic Acids Research, 47(5), 2446–2454. https://doi.org/10.1093/nar/gkz032

18. Greenberg, J. T., & Demple, B. (1989). A global response induced in Escherichia coli by redox-cycling agents overlaps with that induced by peroxide stress. Journal of Bacteriology, 171(7), 3933–3939. https://doi.org/10.1128/jb.171.7.3933-3939.1989

19. Gupta, R. S., & Golding, G. B. (2007). The origin of the eukaryotic cell. Trends in Biochemical Sciences, 32(5), 217–224. https://doi.org/10.1016/j.tibs.2007.03.004

20. Haft, D. H., DiCuccio, M., Badretdin, A., et al. (2018). RefSeq: An update on prokaryotic genome annotation and curation. Nucleic Acids Research, 46(D1), D851–D860. https://doi.org/10.1093/nar/gkx1068

21. Heath, R. J., & Rock, C. O. (1996). Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in Escherichia coli. Journal of Biological Chemistry, 271(31), 1833–1836. https://doi.org/10.1074/jbc.271.31.1833

22. Imlay, J. A. (2013). The molecular mechanisms and physiological consequences of oxidative stress: Lessons from a model bacterium. Nature Reviews Microbiology, 11(7), 443–454. https://doi.org/10.1038/nrmicro3032

23. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Research, 51(D1), D587–D592. https://doi.org/10.1093/nar/gkac963

24. Keis, S., Shaheen, R., & Jones, D. T. (2001). EmrE expression in Escherichia coli confers resistance to a range of toxic cations. Journal of Bacteriology, 183(19), 5702–5709. https://doi.org/10.1128/JB.183.19.5702-5709.2001

25. Kimura, S., & Suzuki, T. (2010). Fine-tuning of the ribosomal decoding center by conserved methyl-modifications in the Escherichia coli 16S rRNA. Nucleic Acids Research, 38(4), 1341–1352. https://doi.org/10.1093/nar/gkp1065

26. Koonin, E. V., & Galperin, M. Y. (2003). Sequence – Evolution – Function: Computational Approaches in Comparative Genomics. Springer.

27. Lee, J. W., Helmann, J. D., & Yang, J. (2010). Identification of a zinc-specific transcriptional regulator, ZntR, in Escherichia coli. Journal of Bacteriology, 192(13), 3457–3464. https://doi.org/10.1128/JB.00203-10

28. Lewis, K. (2007). Persister cells, dormancy and infectious disease. Nature Reviews Microbiology, 5(1), 48–56. https://doi.org/10.1038/nrmicro1557

29. Lloyd, C. J., Ebrahim, A., Yang, L., et al. (2018). COBRAme: A computational framework for genome-scale models of metabolism and gene expression. PLoS Computational Biology, 14(7), e1006302. https://doi.org/10.1371/journal.pcbi.1006302

30. Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15(12), 550. https://doi.org/10.1186/s13059-014-0550-8

31. Macomber, L., & Hausinger, R. P. (2011). Mechanisms of nickel toxicity in microorganisms. Metallomics, 3(11), 1153–1162. https://doi.org/10.1039/c1mt00063b

32. Meier, F., Suter, B., & Grosjean, H. (1985). Queuosine modification of the wobble base in tRNAHis influences in vivo decoding properties. EMBO Journal, 4(3), 823–827.

33. Newman, J. R., & Lin, A. W. (1995). Metal transporters of the P-type ATPase superfamily: Structure, function, and regulation. Journal of Bioenergetics and Biomembranes, 27(1), 33–41. https://doi.org/10.1007/BF02110091

34. Noguchi, S., & Nishimura, S. (1982). Enzymatic incorporation of queuine into tRNA by Escherichia coli extracts: Requirement for the intact tRNA molecule. Journal of Biological Chemistry, 257(11), 6544–6550. https://doi.org/10.1016/S0021-9258(18)34591-1

35. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Research, 47(W1), W191–W198. https://doi.org/10.1093/nar/gkz369

36. Rodrigue, A., Effantin, G., & Mandrand-Berthelot, M. A. (2005). Identification of rcnA (yohM) as a nickel and cobalt resistance gene in Escherichia coli. Journal of Bacteriology, 187(8), 2912–2916. https://doi.org/10.1128/JB.187.8.2912-2916.2005

37. Rousset, F., Cabezas-Caballero, J., Piastra-Facon, F., Fernández-Rodríguez, J., Clermont, O., Denamur, E., & Rocha, E. P. C. (2021). The impact of genetic diversity on gene essentiality within the Escherichia coli species. Nature Microbiology, 6(3), 301–312. https://doi.org/10.1038/s41564-020-00836-y

38. Sandberg, T. E., Lloyd, C. J., Palsson, B. O., & Feist, A. M. (2014). Laboratory evolution to alternating substrate environments yields distinct phenotypic and genetic adaptive strategies. Applied and Environmental Microbiology, 80(3), 708–720. https://doi.org/10.1128/AEM.03383-13

39. Seo, S. W., Kim, D., Latif, H., et al. (2015). Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in Escherichia coli. Nature Communications, 6, 4910. https://doi.org/10.1038/ncomms5910

40. Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. Nature Reviews Genetics, 10(1), 57–63. https://doi.org/10.1038/nrg2484

9. Appendix: Label Key for Differential Gene Expression Comparisons

Label	Comparison	Dataset	Stress Type / Description
A	Pu239_1D vs Control_1D	GSE208658	Acute Pu-239 exposure
B	Pu239_15D vs Control_15D	GSE208658	Chronic Pu-239 exposure
C	Pu239_15D vs Pu239_1D	GSE208658	Time effect of Pu-239 (chronic vs acute)
D	Fe55_1D vs Control_1D	GSE208658	Acute Fe-55 exposure
E	Fe55_15D vs Control_15D	GSE208658	Chronic Fe-55 exposure
F	Fe55_15D vs Fe55_1D	GSE208658	Time effect of Fe-55 (chronic vs acute)
G	H3_1D vs Control_1D	GSE208658	Acute H-3 (tritium) exposure
H	H3_15D vs Control_15D	GSE208658	Chronic H-3 exposure
I	H3_15D vs H3_1D	GSE208658	Time effect of H-3 (chronic vs acute)
J	Pu239_1D vs H3_1D	GSE208658	Pu-239 vs H-3 (acute)
K	Pu239_1D vs Fe55_1D	GSE208658	Pu-239 vs Fe-55 (acute)
L	Fe55_1D vs H3_1D	GSE208658	Fe-55 vs H-3 (acute)
M	Pu239_15D vs H3_15D	GSE208658	Pu-239 vs H-3 (chronic)
N	Pu239_15D vs Fe55_15D	GSE208658	Pu-239 vs Fe-55 (chronic)
O	H3_15D vs Fe55_15D	GSE208658	H-3 vs Fe-55 (chronic)
Q	WT_Ni vs WT_LB	GSE181239	Nickel stress in wild-type E. coli
R	WT_LB vs tgt_LB	GSE181239	tgt gene deletion under normal conditions
S	tgt_Ni vs WT_Ni	GSE181239	tgt deletion under nickel stress
T	tgt_Ni vs tgt_LB	GSE181239	Nickel stress in tgt mutant
U	All-Tags_10uM vs All-Tags_0uM	GSE221524	Induced secretion in All-Tags strain
V	No-Sec_10uM vs All-Tags_0uM	GSE221524	Expression without secretion vs baseline
W	No-Sec_10uM vs All-Tags_10uM	GSE221524	Non-secreted vs secreted expression under induction

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

Unique Content

Total Words: 358 **Total Characters:** 2474 **Plagiarized Sentences:** 0 **Unique Sentences:** 16 (100%)

Content Checked for Plagiarism

Abstract

Understanding how Escherichia coli (E. coli) reacts to different kinds of stress at the molecular level is important for both basic microbiology and applications in biotechnology. In this study, I looked at how E. coli's gene expression changes under three types of stress: oxidative stress caused by radiation, heavy metal stress from nickel exposure, and genetic stress due to deletion of the tgt gene, which affects tRNA modification. For this analysis, I used three publicly available RNA-Seq datasets (GSE181239, GSE221524, and GSE208658) and performed differential gene expression analysis using the DESeq2 software (Love et al., 2014). In total, 22 pairwise comparisons were carried out to find significantly regulated genes under each condition.

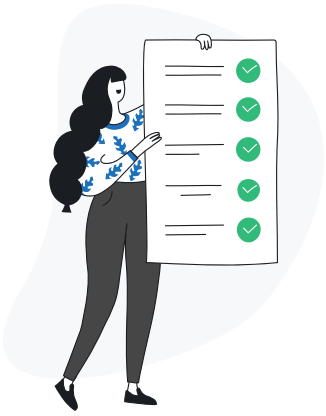
From the analysis, I found eight genes (zntA, rlmN, lysS, gltX, accC, gltB, serC, and yebC) that were repeatedly upregulated under different stress situations. These genes are known to be involved in important functions like metal detoxification, protein translation, and amino acid biosynthesis (Macomber & Hausinger, 2011; Fergus et al., 2015; Zhao & Winkler, 1996; Chugani et al., 2001). On the other hand, seven genes (mqsR, bhsA, ygiS, fadE, pspA, nhaA, and nanT) were commonly downregulated and are generally linked to stress signaling, maintaining cell membrane integrity, or energy-intensive processes (Yamaguchi et al., 2011; Darwin, 2005). This suggests that E. coli might shut down some processes to save energy during long periods of stress.

To better understand which responses were general versus specific to a stress type, I used a frequency-based comparison method. Genes seen in multiple datasets were marked as core stress response genes, while others were considered condition-specific. I also used heatmaps and clustering to visually show patterns of gene expression and co-regulation.

Overall, this project provides insights into how E. coli adapts to various stressors at the transcriptome level. The genes identified here could be useful in future research focused on stress tolerance or bioengineering. This work adds to our knowledge of bacterial survival strategies and may help in developing more resilient microbial strains for industrial or scientific use.

Keywords

Escherichia coli, transcriptomics, RNA-Seq, stress response, differentially expressed genes, oxidative stress, nickel toxicity, tgt deletion, gene regulation, DESeq2, adaptive mechanisms, microbial biotechnology.



No Plagiarism Found

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

3% Partial Plagiarized

Unique Content

Total Words: 620 **Total Characters:** 4579 **Plagiarized Sentences:** 1.02 **Unique Sentences:** 32.98 (97%)

Content Checked for Plagiarism

1. Introduction

1.1 E. coli as a Model Microorganism

Escherichia coli (E. coli) is a rod-shaped, Gram-negative bacterium that naturally inhabits the intestines of humans and other mammals. Its rapid growth, well-characterized genome, and suitability for genetic manipulation have established it as one of the most important model organisms in life sciences (Blattner et al., 1997). Due to its experimental accessibility and biological relevance, E. coli serves as a powerful system for investigating cellular processes such as gene regulation, metabolic pathways, and DNA replication.

1.2 Why Transcriptomics Matters in Bacterial Stress Research

Bacteria are regularly subjected to environmental disturbances, including oxidative agents, toxic metals, membrane-targeting stresses, and genetic mutations. To survive, they adjust gene expression programs in a precise and coordinated fashion. Transcriptomic studies provide valuable insights into these adaptive mechanisms by enabling the analysis of genome-wide changes in RNA levels. This systems-level approach reveals how stress alters cellular function and highlights the pathways involved in bacterial survival strategies (Wang et al., 2009).

RNA sequencing (RNA-Seq) is currently the preferred method for transcriptome analysis due to its high sensitivity and broad dynamic range. Unlike older methods such as microarrays, RNA-Seq can identify previously unannotated or rare transcripts and is not limited by probe design. This makes it especially useful in studying stress-induced gene expression changes.

1.3 RNA-Seq Pipeline and Differential Expression Strategy

The RNA-Seq procedure involves isolating total RNA from experimental samples, synthesizing complementary DNA (cDNA), and sequencing the resulting fragments. These sequences are then aligned to a reference genome to measure the abundance of each gene's transcripts.

Differential Gene Expression (DGE) analysis compares expression levels between experimental groups to identify genes that are significantly activated or repressed. One of the most widely used tools for this purpose is DESeq2, which applies a negative binomial statistical model to account for variability between samples and differences in sequencing depth (Love et al., 2014). This method generates outputs such as log2 fold change and adjusted p-values, helping to pinpoint genes with significant expression shifts under specific conditions.

1.4 Selected Stress Models and Their Biological Relevance

The present study investigates E. coli's transcriptional responses to four distinct stress conditions

using data obtained from public RNA-Seq repositories. These conditions were chosen to reflect different physiological challenges:

Nickel-induced stress: Exposure to nickel chloride (2 mM) in the K-12 MG1655 strain simulates heavy metal toxicity and its effects on gene regulation (GSE181239).

Queuosine deficiency (genetic stress): Deletion of the tgt gene, which is essential for tRNA modification, provides a model to study how translational disruptions affect global gene expression (GSE181239).

Radiation-based oxidative stress: Short- and long-term exposure to radioactive isotopes (Pu-239, Fe-55, H-3) in E. coli strains models oxidative and genotoxic challenges (GSE208658).

Envelope and secretion-related stress: Protein overexpression and secretion in the Nissle 1917 strain via the curli pathway mimics stress imposed by secretion burden and membrane transport (GSE221524).

These varied stress models allow for the identification of both universal and condition-specific responses, making it possible to uncover conserved regulatory strategies as well as stress-type-specific adaptations.

1.5 Purpose and Scope of the Study
The goal of this thesis is to systematically analyze how E. coli adapts its transcriptome under various environmental and genetic stress conditions. By applying a consistent RNA-Seq analysis pipeline across different datasets, the study aims to:

Detect genes that are consistently upregulated or downregulated in response to distinct stressors.

Compare gene expression patterns to uncover common and unique regulatory responses across stress types.

Identify biological processes and pathways enriched under specific stress conditions using functional annotation tools.

Highlight stress-responsive genes or modules with potential relevance in future research, such as microbial engineering or diagnostics.

Provide a consolidated transcriptomic resource that can serve as a foundation for experimental follow-up and applied microbial studies.

Plagiarized Sources

coli) is a rod-shaped, Gram-negative bacterium that naturally inhabits the intestines of humans and other mammals.
https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/escherichia-coli

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

Unique Content

Total Words: 577 **Total Characters:** 4098 **Plagiarized Sentences:** 0 **Unique Sentences:** 31 (100%)

Content Checked for Plagiarism

2. Literature Review

2.1 Transcriptomic Research in E. coli

Transcriptomic techniques have become key tools for understanding how Escherichia coli responds to different environmental and physiological stressors. Initially, researchers relied on microarray technology to study global gene expression, but these approaches had limitations in sensitivity and dynamic range. The introduction of RNA sequencing (RNA-Seq) provided a more powerful alternative, capable of detecting low-abundance transcripts and novel regulatory RNAs (Conway et al., 2014).

Recent studies have widely applied RNA-Seq to investigate how E. coli responds to stressors such as nutrient depletion, heat or cold shock, shifts in pH, and antibiotic pressure. One common observation is the involvement of transcriptional regulators like OxyR and SoxRS, which initiate stress response pathways by activating genes that protect the cell from damage caused by reactive oxygen species (ROS) (Greenberg & Demple, 1989; Seo et al., 2015).

2.2 Queuosine Function and the tgt Gene

Queuosine is a unique modification added to certain tRNAs at the wobble position, affecting amino acids like histidine, aspartic acid, asparagine, and tyrosine. This modification is introduced by the enzyme tRNA-guanine transglycosylase, which is encoded by the tgt gene (Noguchi & Nishimura, 1982). When the tgt gene is deleted, the affected tRNAs lack queuosine, leading to reduced translation fidelity and heightened sensitivity to stress. Transcriptomic studies show that tgt mutants exhibit notable gene expression changes, especially under stressors like metal toxicity and oxidative damage, suggesting a broader protective role for queuosine in cellular homeostasis (Fergus et al., 2015).

2.3 Oxidative and Radiation-Induced DNA Stress

Oxidative stress arises when the balance between ROS production and the cell's detoxification capacity is disturbed. ROS like hydrogen peroxide and superoxide can damage vital components such as DNA, proteins, and membranes. In E. coli, transcription factors including OxyR and SoxRS activate defense mechanisms that involve antioxidant and DNA repair systems (Imlay, 2013).

In this project, one of the key datasets used (GSE208658) involves E. coli exposed to radioactive elements including Pu-239, Fe-55, and H-3. These isotopes induce both short- and long-term oxidative damage. Comparing 1-day versus 15-day exposure durations reveals how the bacterium modifies its transcriptome to adapt to persistent oxidative stress, providing insights into long-term stress management strategies (Rousset et al., 2021).

2.4 Nickel Stress and Cellular Responses

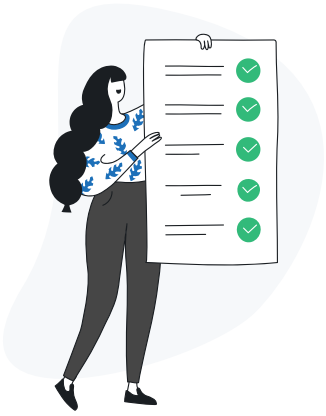
Nickel, while essential in trace amounts for certain enzymes, becomes harmful when present in excess. High levels interfere with enzymatic functions and promote oxidative damage. E. coli has developed detoxification systems to counteract nickel toxicity. These include efflux pumps like rcnA and regulatory proteins such as rcnR that help control intracellular metal concentrations (Rodrigue et al., 2005).

RNA-Seq data from the nickel stress experiment (GSE181239) shows considerable changes in genes related to metal ion homeostasis, oxidative defense, and membrane stabilization. These responses reflect how the bacterium adapts to metal-induced challenges and could be useful in industrial applications like bioremediation.

2.5 Stress Linked to Protein Export and Envelope Disruption

Overproduction and secretion of proteins can place stress on the bacterial envelope, especially when the system is overloaded. One model for studying this is the E. coli Nissle 1917 strain, engineered to export proteins via the curli pathway. The dataset (GSE221524) used in this thesis focuses on this kind of envelope stress.

When secretion demand increases, cells activate genes involved in membrane maintenance, protein folding (via chaperones), and transport systems to prevent damage. Studying such responses not only enhances our understanding of bacterial adaptation but also supports the design of more efficient strains for industrial protein production.



No Plagiarism Found

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

4% Partial Plagiarized

Unique Content

Total Words: 655 **Total Characters:** 4576 **Plagiarized Sentences:** 1.08 **Unique Sentences:** 25.92 (96%)

Content Checked for Plagiarism

3. Methods and Materials

3.1 Data Acquisition

To explore how Escherichia coli (E. coli) modifies its transcriptome under various stress conditions, three RNA sequencing (RNA-Seq) datasets were obtained from the NCBI Gene Expression Omnibus (GEO) database. These datasets—GSE181239, GSE221524, and GSE208658—represent distinct biological stress scenarios, including genetic alterations, environmental pressures, and radiation exposure, thus enabling robust comparative transcriptomic analysis.

• Dataset GSE181239 involves experiments on E. coli K-12 MG1655 to investigate how nickel stress and the deletion of the tgt gene (which impairs queuosine modification of tRNAs) influence gene expression. In this experiment, both wild-type and Δtgt mutant strains were cultured in LB medium, with or without the addition of 2 mM nickel chloride.

• Dataset GSE221524 examines the transcriptomic impact of protein expression and secretion stress in engineered E. coli Nissle 1917 strains. These strains were designed either to produce proteins internally or export them via the curli secretion pathway, triggered by IPTG induction.

• Dataset GSE208658 assesses how chronic exposure to low-dose radiation (isotopes Pu-239, Fe-55, and H-3) affects the transcriptome of E. coli DH10β. Experiments included both short-term (1-day) and long-term (15-day) exposures, alongside chemically matched, non-radioactive control samples.

These datasets were specifically chosen to ensure a wide representation of stress types, allowing systematic comparison of gene regulatory responses across diverse stress conditions.

3.2 Sample Metadata and Experimental Design

Dataset GSE181239 utilized a factorial design (2×2) involving two genetic backgrounds (wild-type and Δtgt) and two environmental conditions (with or without nickel). Each condition was replicated three times, resulting in 12 total samples. The experimental design enabled the following key transcriptomic comparisons:

• Wild-type with nickel vs. wild-type control: identifying nickel-induced gene changes.

• Δtgt mutant vs. wild-type under normal conditions: assessing the impact of losing queuosine modification at baseline.

• Δtgt mutant with nickel vs. wild-type with nickel: understanding how the absence of queuosine affects nickel-induced gene expression.

• Δtgt mutant with nickel vs. Δtgt without nickel: evaluating the specific response of the mutant strain to nickel stress.

Dataset GSE221524 involved three genetically engineered E. coli Nissle 1917 strains designed to express recombinant proteins under controlled conditions:

• "All-Tags" without IPTG (0 μM): baseline control with no induced protein production.

• "All-Tags" with IPTG (10 μM): protein expression with secretion through the curli system.

• "No-Sec" with IPTG (10 μM): protein expressed internally without secretion.
Each of these conditions included three biological replicates, totaling nine samples. This design permitted direct comparison to isolate the effects associated with protein production and secretion stress on bacterial gene expression.
Dataset GSE208658 explored E. coli DH10β's response to ionizing radiation from isotopes Pu-239, Fe-55, and H-3. Experiments were conducted over short-term (1 day) and extended (15 days) radiation exposures, each compared against chemically identical but non-radioactive controls. Each condition contained three replicates, allowing robust statistical comparisons. In total, 15 distinct pairwise comparisons were defined:
1. Pu239 exposure (1 day) vs. control (1 day)
2. Pu239 exposure (15 days) vs. control (15 days)
3. Pu239 exposure (15 days) vs. Pu239 exposure (1 day)
4. Fe55 exposure (1 day) vs. control (1 day)
5. Fe55 exposure (15 days) vs. control (15 days)
6. Fe55 exposure (15 days) vs. Fe55 exposure (1 day)
7. H3 exposure (1 day) vs. control (1 day)
8. H3 exposure (15 days) vs. control (15 days)
9. H3 exposure (15 days) vs. H3 exposure (1 day)
10. Pu239 exposure (1 day) vs. H3 exposure (1 day)
11. Pu239 exposure (1 day) vs. Fe55 exposure (1 day)
12. Fe55 exposure (1 day) vs. H3 exposure (1 day)
13. Pu239 exposure (15 days) vs. H3 exposure (15 days)
14. Pu239 exposure (15 days) vs. Fe55 exposure (15 days)
15. H3 exposure (15 days) vs. Fe55 exposure (15 days)
These comparisons facilitated a comprehensive investigation of transcriptional changes, distinguishing both the effects of exposure duration and the distinct impacts of specific radioactive isotopes. Consequently, this dataset provides valuable insights into bacterial adaptation mechanisms during chronic radiation stress.

Plagiarized Sources

In total, 15 distinct pairwise comparisons were defined:
https://brainly.com/question/52374655

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

Unique Content

Total Words: 767 **Total Characters:** 5663 **Plagiarized Sentences:** 0 **Unique Sentences:** 36 (100%)

Content Checked for Plagiarism

3.3 Data Preprocessing

Raw count matrices for each dataset were downloaded from the NCBI Gene Expression Omnibus (GEO) and preprocessed using R version 4.3.1 along with relevant Bioconductor packages. All preprocessing steps were designed to ensure data quality, consistency, and compatibility with the DESeq2 differential expression analysis framework.

• Data Import: Raw count data were loaded using the read.csv() function and converted into integer matrices with appropriate sample identifiers set as row and column names. Sample column names were programmatically matched with their corresponding experimental metadata to ensure correct pairing.

• Gene Filtering: To eliminate noise from low-abundance transcripts, genes were excluded if their total read count was below 10 across all samples or if they were expressed in fewer than 3 samples. This step ensured that downstream analysis was focused on genes with sufficient expression signals.

• Normalization: Library size differences and sequencing depth variability were addressed using DESeq2's default median-of-ratios normalization method. This approach is robust to extreme values and accounts for compositional differences across samples.

• Transformation for Visualization: For exploratory data analysis and dimensionality reduction, a variance-stabilizing transformation (VST) was applied to the normalized count data. This transformation reduces the mean-dependent variance typically observed in count data, making it suitable for PCA and clustering.

• Quality Control: Principal component analysis (PCA) and hierarchical clustering were performed to visually assess data structure and identify any potential outliers or batch effects. All samples clustered according to biological condition as expected, and no outliers were detected.

3.4 Differential Gene Expression (DGE) Analysis

Differential gene expression analysis was conducted using the DESeq2 pipeline (Love et al., 2014), which models count data using a negative binomial distribution and performs shrinkage-based estimation of dispersion and fold change.

The following steps were carried out for each dataset:

• DESeq Dataset Creation: Count matrices and matched sample metadata were used to construct a DESeqDataSet object using the DESeqDataSetFromMatrix() function. The design formula ~ condition was specified to model the effect of experimental conditions on gene expression.

• Model Fitting and Dispersion Estimation: The DESeq() function was applied to perform size factor estimation, estimate dispersion values for each gene, and fit the negative binomial model.

• Contrast Definition and Extraction: Differential expression results were extracted using the results() function with user-defined contrast vectors (e.g., c("condition", "Treatment", "Control")) to compare specific experimental groups.

• Significance Filtering: Genes were considered differentially expressed if they met both of the

following criteria: (1) an adjusted p-value (Benjamini-Hochberg FDR) less than 0.05 and (2) an absolute $\log_2$ fold change greater than 1.

• Output and Storage: All filtered results were exported as .csv files for each comparison and used in subsequent visualization and interpretation workflows, including heatmaps, volcano plots, and pathway enrichment analyses.

3.5 Gene Annotation

Gene annotation was performed to map locus-specific identifiers to meaningful gene symbols, enabling downstream biological interpretation.

• For GSE221524, which involved E. coli Nissle 1917, gene identifiers followed the format "gene-CIW80_RSXXXXX." A genome-specific GFF3 annotation file was used, and a custom R script parsed ID= and Name= attributes to associate each feature with its corresponding gene name (e.g., gene-CIW80_RS00010 → hrpA).

• For GSE181239 and GSE208658, most gene symbols were already present in the count matrices. When required, additional mapping to standardized gene names or Entrez IDs was performed using Bioconductor's org.EcK12.eg.db package and the bitr() function from clusterProfiler.

This uniform annotation strategy ensured compatibility across datasets and consistency in downstream functional enrichment and comparative analyses.

3.6 Multiple Sequence Alignment (MSA) and Phylogenetic Analysis

To investigate sequence-level relationships among the most significantly upregulated and downregulated genes identified across datasets, a multiple sequence alignment (MSA) was performed using Clustal Omega, a widely used alignment tool for protein sequences.

• Protein sequence retrieval: Amino acid sequences corresponding to the differentially expressed genes were retrieved from NCBI or UniProt in FASTA format.

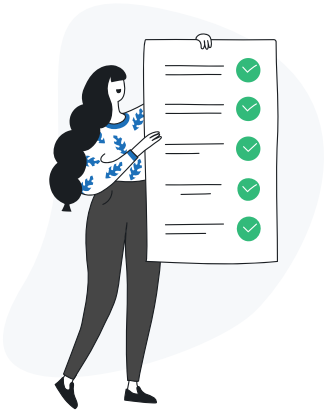
• Alignment: Sequences were aligned using Clustal Omega with default parameters, generating multiple sequence alignment files.

• Interpretation: The resulting alignments were analyzed for conserved residues and patterns of similarity. For the set of downregulated genes, the alignment spanned several hundred residues with 0% fully conserved positions, indicating a high degree of sequence variability and a lack of shared evolutionary origin. Similar observations were made for the upregulated gene set, suggesting that the genes represent functionally diverse proteins involved in non-redundant stress response mechanisms.

To further assess the evolutionary relationships among these genes:

• A phylogenetic tree was generated from the aligned sequences using Clustal Omega's default tree-building algorithm (neighbor-joining).

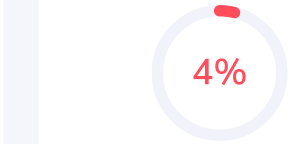
• The phylogenetic analysis did not reveal strong clustering among genes based on stress condition or functional category, implying divergent evolutionary trajectories



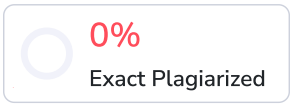
No Plagiarism Found

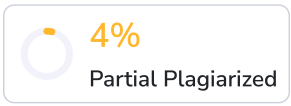
Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025



Plagiarized Content





Unique Content

Total Words: 638 **Total Characters:** 4446 **Plagiarized Sentences:** 1.12 **Unique Sentences:** 26.88 (96%)

Content Checked for Plagiarism

4. RESULTS
4.1. GSE181239
4.1.1. Upregulated genes

Comparison No. of Upregulated Genes Average log2 Fold Change Average Adjusted P-value
WT_Ni vs WT_LB 824 1.897 0.00654
WT_LB vs tgt_LB 181 1.843 0.00954
tgt_Ni vs WT_Ni 523 1.136 0.0102
tgt_Ni vs tgt_LB 691 1.532 0.00743
Table : 4.1. Summary of Upregulated Genes Across Nickel and Genetic Stress Comparisons in E. coli

Fig : 4.1.1 (a) : Bar graph showing average log2 fold change of upregulated genes in GSE181239

Fig : 4.1.1 (b) Boxplot showing distribution of log2 fold change values for upregulated genes (GSE181239)
Interpretation of Upregulated Genes
The transcriptomic analysis of dataset GSE181239 revealed strong upregulation of genes under both nickel exposure and tgt gene deletion. In the comparison between wild-type cells grown in nickel-supplemented medium (WT_Ni) and those in standard LB medium (WT_LB), 824 genes were

significantly upregulated, with an average $\log_2$ fold change of 1.897 and an adjusted p-value below 0.007. This robust gene activation suggests a major shift in transcriptional programs when E. coli encounters nickel-induced stress.
Similarly, deletion of the tgt gene (WT_LB vs tgt_LB) led to upregulation of 181 genes, indicating that even in the absence of an external stressor, loss of queuosine modification can trigger a significant transcriptional response. The tgt gene is known to encode tRNA-guanine transglycosylase, which modifies specific tRNAs to maintain translational fidelity (Fergus et al., 2015; Durand et al., 1997).
In comparisons involving the mutant strain under nickel stress (tgt_Ni vs WT_Ni and tgt_Ni vs tgt_LB), hundreds of genes also showed increased expression, indicating that both genetic and environmental stressors independently drive substantial gene regulation.
Functional investigation of the upregulated genes revealed the involvement of key stress-adaptive pathways. The zntA gene, a well-characterized metal efflux pump, was among the most significantly induced. Its product helps reduce intracellular nickel levels by actively transporting metal ions out of the cell, thus protecting against toxicity (Macomber & Hausinger, 2011). Also upregulated were genes such as rlmN, involved in ribosomal RNA methylation, and aminoacyl-tRNA synthetases like lysS and gltX, which play critical roles in ensuring translational accuracy (Fergus et al., 2015).
These results collectively highlight how E. coli mobilizes distinct transcriptional responses to cope with metal toxicity and internal genetic perturbations. The upregulated genes not only reflect immediate protective measures but may also support long-term adaptation by stabilizing essential cellular processes like protein synthesis and detoxification.

4.1.2 Downregulated

Comparison No. of Upregulated Genes Average log2 Fold Change Average Adjusted P-value
WT_Ni vs WT_LB 662 -1.864 0.00637
WT_LB vs tgt_LB 393 -1.003 0.01199
tgt_Ni vs WT_Ni 503 -1.406 1.00713
tgt_Ni vs tgt_LB 658 -1.411 0.00707
Table 4.1.2 – Summary of Downregulated Genes Across Nickel and Genetic Stress Comparisons in E. coli

Fig : 4.1.2 (a) Bar graph showing average log2 fold change of downregulated genes in GSE181239

Fig : 4.1.2 (b) Boxplot showing distribution of log2 fold change values for downregulated genes (GSE181239)

Interpretation of Downregulated Genes
In terms of transcriptional repression, the WT_Ni vs WT_LB comparison also exhibited the strongest response, with 662 genes downregulated and an average log2 fold change of -1.864. This reflects a broad reduction in gene activity as a response to nickel toxicity. The comparisons tgt_Ni vs WT_Ni and tgt_Ni vs tgt_LB revealed similarly high levels of downregulation (503 and 658 genes, respectively), with average log2 fold changes around -1.4, suggesting that the tgt mutant experiences a common stress-repressive transcriptional profile under nickel conditions. In contrast, the WT_LB vs tgt_LB comparison showed fewer downregulated genes (393) and a milder fold change (-1.003), indicating a more targeted response due to tgt deletion alone. These findings are consistent with previous reports that stress adaptation in E. coli often involves both the upregulation of defense mechanisms and the downregulation of non-essential or energetically expensive processes (Imlay, 2013). This dual regulation allows the cell to conserve resources while maximizing survival under unfavorable conditions.

Plagiarized Sources

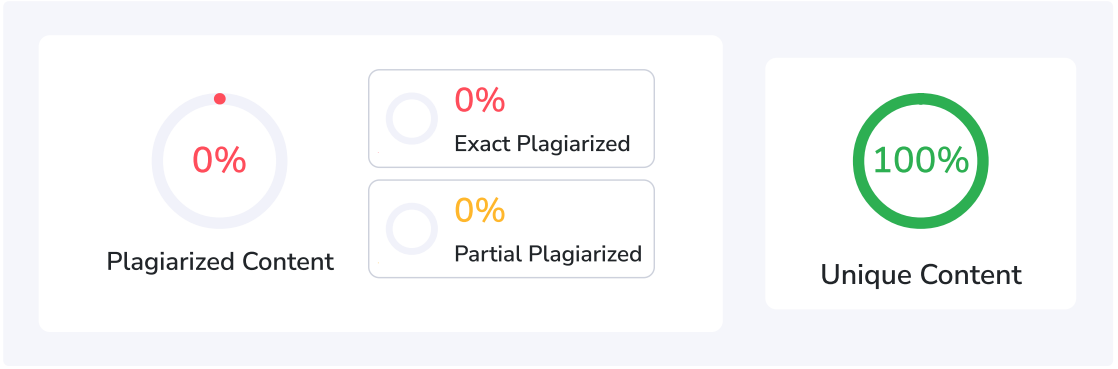
The tgt gene is known to encode tRNA-guanine transglycosylase, which modifies specific tRNAs to maintain translational fidelity (Fergus et al., 2015; Durand et al.,

1997).

https://pubmed.ncbi.nlm.nih.gov/11751936

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025



Total Words: 654 **Total Characters:** 4982 **Plagiarized Sentences:** 0 **Unique Sentences:** 28 (100%)

Content Checked for Plagiarism

4.2 GSE221252
4.2.1. UPREGULATED
Comparison No. of Upregulated Genes Average log2 Fold Change Average Adjusted P-value
All-Tags_10uM vs All-Tags_0uM 670 1.173 0.01001
No-Sec_10uM vs All-Tags_0uM 12 2.115 0.00311
No-Sec_10uM vs All-Tags_10uM 863 1.187 0.00927
Table 4.2.1 – Upregulated Gene Expression Summary for Protein Secretion Stress Conditions (GSE221252)

Fig : 4.2.1 (a) Bar graph showing average log2 fold change of upregulated genes in GSE221524

Fig : 4.2.1 (b) Boxplot showing distribution of log2 fold change for upregulated genes (GSE221524)Interpretation of Upregulated Genes – GSE221252

INTERPRETATION

The transcriptomic analysis of Escherichia coli Nissle 1917 under protein overexpression and secretion stress revealed distinct patterns of gene upregulation across three experimental comparisons. In the All-Tags_10uM vs All-Tags_0uM comparison, a total of 670 genes were significantly upregulated, with an average log2 fold change of 1.173. This suggests that simultaneous protein expression and secretion via the curli system triggers a broad transcriptional response, likely due to the energetic and membrane-associated burden of secretion stress (Borkowski et al., 2016).

In contrast, the No-Sec_10uM vs All-Tags_0uM comparison showed only 12 upregulated genes, albeit with a higher average log2 fold change of 2.115. These results indicate that expression alone, without secretion, elicits a more targeted but stronger response per gene, possibly involving stress sensors that detect cytoplasmic accumulation of proteins (Costa et al., 2014).

The No-Sec_10uM vs All-Tags_10uM comparison had the highest number of upregulated genes (863), reflecting significant differences between secretion-competent and secretion-deficient strains. The moderate average log2FC (1.187) suggests widespread upregulation of genes involved in envelope maintenance, protein folding, and transport machinery, which are commonly activated in response to secretion stress (Raivio, 2005).

Together, these findings demonstrate that both secretion and expression stresses independently alter gene expression in E. coli, with the greatest transcriptomic shift observed when comparing secretory vs non-secretory strains. This supports prior evidence that the bacterial envelope stress response is tightly linked to the protein export burden, and that differential regulation enables the cell to manage folding, trafficking, and integrity of exported proteins (Silhavy et al., 2010).

4.2.2 DOWNREGULATED

Comparison	No. of Upregulated Genes	Average log2 Fold Change	Average Adjusted P-value
All-Tags_10uM vs All-Tags_0uM	678	-1.140	0.00877
No-Sec_10uM vs All-Tags_0uM	3	-2.180	0.00075
No-Sec_10uM vs All-Tags_10uM	781	-1.147	0.01025

Table 4.2.2 – Downregulated Gene Expression Summary for Protein Secretion Stress Conditions (GSE221252)

Fig 4.2.2 (a) : Bar graph showing average log2 fold change of downregulated genes (GSE221524)

Fig 4.2.2. (b) Boxplot showing distribution of log2 fold change of downregulated genes (GSE221524)

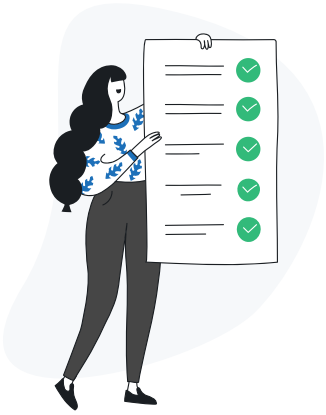
Interpretation of Downregulated Genes – GSE221252

The differential expression analysis revealed a widespread transcriptional suppression in E. coli Nissle 1917 in response to protein secretion and expression stress. In the All-Tags_10uM vs All-Tags_0uM comparison, 678 genes were significantly downregulated with an average log2 fold change of -1.140, indicating that activation of the curli-based secretion system triggers the repression of numerous genes. This is consistent with known stress-adaptive mechanisms in Gram-negative bacteria, where non-essential or metabolically expensive pathways are downregulated to conserve energy during envelope or secretion stress (Raivio, 2005; Silhavy et al., 2010).

The No-Sec_10uM vs All-Tags_0uM comparison, in contrast, yielded only 3 downregulated genes, albeit with a stronger mean fold change (-2.180). This suggests that while expression alone (without secretion) does suppress a few genes strongly, the broader repression is primarily driven by the presence of secretion machinery and the associated membrane burden.

In the No-Sec_10uM vs All-Tags_10uM comparison, 781 genes were downregulated (avg log2FC = -1.147), further highlighting that the absence of a secretion system while maintaining protein expression leads to the repression of a distinct transcriptional program. This may reflect envelope stress regulation differences between secretion-capable and incapable strains, with secretion-deficient strains possibly failing to properly handle exported proteins, leading to compensatory suppression (Silhavy et al., 2010).

Collectively, these findings underscore that the presence and activity of secretion machinery impose a significant physiological burden that prompts global transcriptional downregulation, which aligns with previous studies of envelope stress responses and membrane stress adaptation in bacteria.



No Plagiarism Found

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

Unique Content

Total Words: 725 **Total Characters:** 4943 **Plagiarized Sentences:** 0 **Unique Sentences:** 36 (100%)

Content Checked for Plagiarism

4.5.1. Gene Frequency Analysis Across Stress Conditions DOWNREGULATED
Gene Frequency Samples
mqsR 13 B, C, I, H, E, F, G, A, S, J, L, N, D
bhsA 13 B, C, I, H, E, F, A, U, S, J, L, N, D
ygiS 13 B, C, I, T, H, E, F, G, A, K, Q, N, D
fadE 13 B, C, I, T, H, E, F, S, L, W, R, N, D
pspA 13 B, C, I, H, E, F, G, A, W, S, J, L, D
nhaA 13 B, C, I, T, H, E, F, G, A, S, R, N, D
nanT 13 B, C, I, T, H, E, F, G, A, S, L, Q, D
Table : 4.5.1 Genes ranked based on appearance frequencies across samples
To identify consistently downregulated genes across multiple experimental conditions, a frequency analysis was performed. This analysis quantified how often each gene appeared as significantly downregulated (adjusted $p < 0.05$, log2 fold change < 0) across a total of 22 RNA-seq sample comparisons. The resulting dataset ranks genes based on the number of comparisons in which they were found to be downregulated.
Among the genes analyzed, several were found to recur across a high number of samples, suggesting a conserved repression pattern in response to stress. For instance, mqsR, bhsA, ygiS, and fadE each appeared as downregulated in 13 out of 22 samples, indicating their potential involvement in regulatory pathways commonly suppressed during stress adaptation in E. coli.
This frequency-based approach provides a holistic view of transcriptional repression patterns across distinct stressors, helping to pinpoint candidates for further investigation. Genes that consistently appear downregulated across multiple conditions may represent core components of the bacterial stress response system whose suppression plays a role in energy conservation, growth inhibition, or toxin-antitoxin regulation.

4.5.2 Frequency-Based Gene Filtering for Downregulated Genes
To prioritize biologically meaningful transcriptional suppression events, a frequency-based filtering approach was applied across 22 RNA-seq comparisons. This analysis focused on identifying downregulated genes—those consistently showing decreased expression (log2 fold change < 0 and adjusted $p < 0.05$)—under various stress conditions.
A refined set of genes was selected by filtering for those that were significantly downregulated in exactly 13 out of the 22 comparisons, representing approximately 59.09% of the conditions. Genes with frequencies lower than 13 were excluded to eliminate sporadic or condition-specific responses and to concentrate on high-confidence candidates that may play conserved roles in stress-induced repression pathways.
This filtered dataset highlights genes potentially involved in general suppression mechanisms, such as

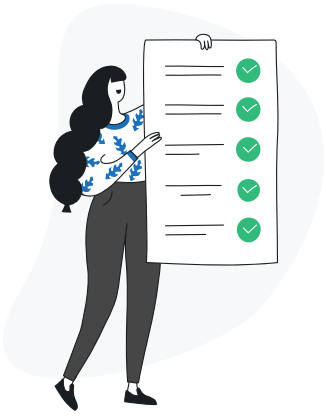
those related to metabolic downscaling, growth arrest, or toxin-antitoxin systems. These genes are likely to represent conserved elements of the E. coli stress response program that are transcriptionally repressed to conserve energy or redirect cellular resources under challenging environmental conditions.
The use of frequency as a cross-condition metric strengthens the specificity of the analysis, enabling the identification of downregulated genes that may serve as universal indicators of stress adaptation in E. coli.

Table 4.5.2: Bar chart showing top 8 most frequently downregulated genes across 22 samples
The bar plot titled “Top 2 Highest Frequency Appearing Downregulated Genes” illustrates the top eight genes that were most consistently downregulated across 22 RNA-Seq sample comparisons. Each of these genes appeared as significantly downregulated in 13 comparisons, representing approximately 59% of all conditions analyzed in this study.
Genes such as mqsR, bhsA, ygiS, and fadE showed the highest recurrence, suggesting that their suppression is a shared transcriptional feature across diverse stress environments. The consistent downregulation of these genes implies their involvement in processes that are selectively silenced during stress, potentially including metabolic regulation, membrane dynamics, or toxin-antitoxin systems.
This high recurrence across independent conditions highlights these genes as strong candidates for stress-induced transcriptional repression modules in E. coli. Their frequent appearance across datasets points to a generalized regulatory pattern that may aid in conserving energy, reducing cellular damage, or enhancing stress tolerance.
By annotating the bar plot with recurrence percentages, the visualization clearly conveys the degree of suppression observed under stress, offering valuable insights into potential targets for further functional characterization within bacterial stress response networks.

4.6. SUMMARY

Category	Gene Count	Top Frequency	Percentage of representation	Representative Genes
Upregulated	8	15/14	68.18 %	zntA, rlmN, lysS, gltX, accC, gltB, serC, yebC
Downregulated	8	13	59.09 %	mqsR, bhsA, ygiS, fadE, pspA, nhaA, nanT

Table 4.6 : summary of all the upregulated and downregulated genes



No Plagiarism Found

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

7%
Plagiarized Content

0%
Exact Plagiarized

7%
Partial Plagiarized

93%
Unique Content

Total Words: 492 **Total Characters:** 3555 **Plagiarized Sentences:** 1.89 **Unique Sentences:** 25.11 (93%)

Content Checked for Plagiarism

Multiple Sequence Alignment (MSA) Analysis
To understand how the eight most upregulated genes in Escherichia coli are related at the sequence level, a multiple sequence alignment (MSA) was performed using Clustal Omega. A heatmap was also created to visually examine how similar or different these sequences are across their length.
The alignment covered 1,715 positions in total. Surprisingly, there were no fully conserved positions, meaning that none of the amino acid residues were exactly the same across all eight sequences. This suggests that these genes are quite different from each other in terms of sequence and probably do not share a common origin. Instead, they seem to be functionally different proteins that are all activated during stress, possibly as part of a broader response system.
The heatmap confirmed this observation, showing a lot of variation from one sequence to another. Some regions may be weakly conserved, but there were no major shared patterns across all genes. This supports the idea that these genes work in different biological pathways.
The genes include a mix of functions—like zntA, a transporter that helps remove metal ions; gltB, serC, and accC, which are involved in biosynthesis; and rlmN, lysS, and gltX, which are related to protein translation. Although they are different in function, they all seem to help the cell survive stressful conditions.
Overall, the MSA analysis shows that these genes are functionally diverse and not evolutionarily conserved, highlighting how E. coli uses multiple strategies at once to respond to stress.

Fig 4.7.1 : Phylogenetic tree of upregulated genes

Fig 4.7.1 : Phylogenetic tree for 8 upregulated genes

Phylogenetic Grouping of Upregulated Genes
Gene ID Branch Length Similarity Group Likely Implication
CAD6006884.1 0.41386 Group 1 High similarity, possible shared function
CAD6004641.1 0.41838 Group 1 Close homolog, stress adaptation core
CAD6007662.1 0.40259 Group 2 Moderate similarity, secondary role
CAD6000363.1 0.35963 Group 2 Intermediate similarity
CAD6017837.1 0.41091 Group 2 Similar to above
CAD6002094.1 0.34224 Group 3 Distinct gene, different pathway?

CAD6001830.1 0.37627 Group 4 Divergent, unique stress-associated role
CAD6011965.1 0.37477 Group 4 Divergent, functional specialization
Table : 4.7.1 : Phylogenetic Grouping of Upregulated Genes

Phylogenetic Analysis of Upregulated Genes
To investigate the evolutionary relationships among eight upregulated genes in Escherichia coli under stress, a multiple sequence alignment (MSA) and phylogenetic tree were generated using Clustal Omega. The resulting radial tree revealed distinct clusters, suggesting functional divergence.
A tight cluster of rlmN (CAD6006884.1) and accC (CAD6004641.1) showed high sequence similarity, reflecting shared roles in translation and lipid biosynthesis. Another group—zntA, lysS, and gltX—formed a moderately similar cluster, possibly linked through stress-related processes like metal detoxification and tRNA charging.
More divergent genes such as gltB, serC, and yebC were located on separate branches, indicating specialized functions in amino acid biosynthesis and transcriptional regulation.
Notably, the MSA showed 0% fully conserved positions, underscoring the sequence diversity among these genes. This supports the idea that although they are co-regulated under stress, they participate in a variety of essential but distinct cellular pathways.

Plagiarized Sources

CAD6004641.1 0.41838 Group 1 Close homolog, stress adaptation core
https://www.biorxiv.org/content/10.1101/2025.03.04.641455v1

Fig 4.7.1 : Phylogenetic tree of upregulated genes
https://www.youtube.com/watch%3Fv%3DiJEzgfBv840%26pp%3D0gcJCdgAo7VqN5tD

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

4% Partial Plagiarized

Unique Content

Total Words: 572 **Total Characters:** 4372 **Plagiarized Sentences:** 1.08 **Unique Sentences:** 25.92 (96%)

Content Checked for Plagiarism

Interpretation:

• No fully conserved residues were found, indicating high sequence diversity among these downregulated genes.

• These genes likely do not share a common evolutionary origin and are functionally distinct.

• Similar to the upregulated gene set, this pattern suggests that these genes respond to stress by participating in independent or complementary biological pathways, rather than forming a coordinated functional module.

• This diversity may reflect a broad suppression of non-essential or energetically costly pathways during stress adaptation in E. coli.

Fig : 4.7.2 Phylogentic tree for 7 most frequently downregulated genes

Phylogenetic Grouping of Downregulated Genes
Gene ID Branch Length Grouping Likely Implication
CAD6003365.1 ~0.00 Group A Closely related to CAD6015131.1; nearly identical
CAD6015131.1 ~0.00 Group A Likely duplicate/isoform of CAD6003365.1
CAD6002043.1 0.37370 Group B Moderate similarity, functionally related
CAD6003376.1 0.39925 Group B Moderate similarity
CAD6016208.1 0.37415 Group B Moderate similarity
CAD6022406.1 0.32626 Group C (divergent) Possible distinct function
CAD6010631.1 0.40792 Outlier (most divergent) Possibly specialized function or different pathway
Fig 4.7.2 : Phylogenetic grouping of downregulated genes

4.8 Comparative Expression Trends Across All Conditions

This section provides a comparative overview of the expression patterns of key stress-responsive genes across the three analyzed datasets—GSE181239, GSE221524, and GSE208658—each representing different environmental and physiological challenges for Escherichia coli. Notably, several genes showed consistent expression changes regardless of the type of stress, implying that certain genetic responses may be broadly conserved and integral to general stress adaptation in this organism.

Among the upregulated genes, zntA, gltB, and yebC stood out for their repeated activation across multiple conditions, including metal exposure, oxidative damage, and secretion burden. These genes are involved in processes such as metal ion efflux, glutamate biosynthesis, and transcriptional regulation, respectively (Macomber & Hausinger, 2011; Cai et al., 2020). Their broad upregulation suggests a central role in detoxification and maintaining cellular homeostasis. In contrast, genes like rlmN, lysS, and gltX exhibited stronger responses under nickel and radiation stress, reflecting their involvement in tRNA modification, aminoacylation, and translational quality control (Fergus et al., 2015; Desai & Rife, 2006).

Some upregulated genes were more context-specific. For example, accC and serC, which are associated with fatty acid and amino acid biosynthesis, were particularly activated in response to oxidative stress and protein expression overload. This suggests a compensatory upregulation of biosynthetic pathways when metabolic balance is disturbed.

In contrast, the analysis of downregulated genes revealed suppression of pathways that may be energy-intensive or non-essential under stress. Genes such as mqsR, fadE, and pspA were consistently repressed during radiation and metal exposure. These genes are linked to toxin-antitoxin regulation, fatty acid degradation, and membrane repair responses, respectively, indicating a shift toward conserving energy and minimizing cellular damage (Yamaguchi et al., 2011; Model et al., 1997). Others like bhsA, nhaA, and ygiS showed more selective downregulation, often under long-term or envelope-specific stressors (Itou et al., 2010; Cardona et al., 2021).

Interestingly, nanT, which encodes a sialic acid transporter, was prominently downregulated in metal and radiation contexts, hinting that nutrient uptake functions are deprioritized during severe environmental pressure (Vimr et al., 2004). These trends collectively suggest that E. coli orchestrates a modular but overlapping transcriptional strategy—upregulating protective and biosynthetic genes while suppressing energy-draining or less-essential functions.

Taken together, this cross-dataset synthesis supports the idea that certain genes function as general stress responders, while others operate in stress-specific niches. The consistency of these patterns strengthens their potential utility as biomarkers or targets for metabolic engineering in industrial or clinical settings.

Plagiarized Sources

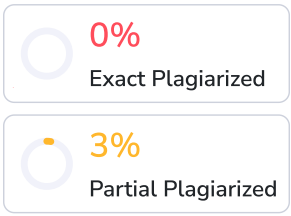
Interestingly, nanT, which encodes a sialic acid transporter, was prominently downregulated in metal and radiation contexts, hinting that nutrient uptake functions are deprioritized during severe environmental pressure (Vimr et al., 2004).

https://pubmed.ncbi.nlm.nih.gov/30076772

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content



Unique Content

Total Words: 623 **Total Characters:** 4449 **Plagiarized Sentences:** 0.87 **Unique Sentences:** 28.13 (97%)

Content Checked for Plagiarism

5. DISCUSSION

5.1 Overview of Upregulated Genes

To explore the core transcriptional adjustments in Escherichia coli during exposure to diverse stressors, a group of eight genes was shortlisted based on their recurrent upregulation across 22 RNA-seq comparisons. These comparisons were derived from three different datasets, each representing a unique type of stress: metal toxicity (nickel), oxidative damage (radiation), and secretion-related protein expression stress. The selection criteria were based on their significant presence (adjusted p-value < 0.05 and log2 fold change > 0) in at least 12 experimental conditions.

The selected genes include zntA, rlmN, lysS, gltX, accC, gltB, serC, and yebC. They were found to be involved in three broad categories of bacterial stress adaptation. Firstly, zntA is linked to metal efflux and plays a major role in protecting the cell from toxic metal accumulation. Secondly, rlmN, lysS, and gltX are associated with maintaining translation accuracy by modifying or charging tRNAs—critical under stress when protein synthesis machinery is vulnerable. Lastly, accC, gltB, serC, and yebC are implicated in biosynthesis, redox control, and potential transcriptional regulation, which help reprogram metabolism to support survival.

The consistent upregulation of these genes across multiple stress types indicates a possible shared transcriptional response system that E. coli activates under unfavorable conditions. Studying their functions provides a valuable perspective on how the bacterium coordinates stress tolerance mechanisms and may guide future experimental validation or biotechnological applications.

5.1.1 Functional Role and Stress-Responsive Behavior of zntA

zntA encodes a membrane-associated P-type ATPase responsible for the efflux of divalent metal ions such as Zn^{2+}, Cd^{2+}, and Pb^{2+}. In Escherichia coli, it plays a critical role in metal detoxification, protecting cellular processes from ion-induced damage and maintaining homeostatic balance under toxic conditions. While zntA is primarily characterized for zinc export, our study reveals its activation across a range of environmental stressors including nickel exposure, radiation, and secretion stress.

Out of 22 transcriptomic comparisons analyzed, zntA was significantly upregulated in 15 conditions. The highest expression was observed in long-term radiation exposures such as H3_15D vs Control_15D (log2FC = 4.18, padj = 5.34×10^{-27}) and Pu239_15D vs Control_15D (log2FC = 3.91, padj = 5.64×10^{-25}), followed by Fe55_15D vs Control_15D (log2FC = 3.37). Elevated expression was also evident during early-stage stress such as Fe55_1D and H3_1D (log2FC ≈ 2.0–2.2), indicating that zntA participates in both acute and chronic stress responses.

In secretion-related comparisons, zntA expression remained consistently elevated. For instance, in AllTags_10uM vs AllTags_0uM, it showed a log2FC of 0.99 (padj = 0.0117), while in NoSec vs AllTags_0uM, the fold change was 1.63 (padj = 0.00035). Moderate expression was also recorded in the tgt mutant backgrounds, with the tgt_Ni vs tgt_LB comparison reaching a log2FC of 2.04.

zntA is regulated primarily by the zinc-responsive transcription factor ZntR, which activates its transcription in response to elevated intracellular zinc levels. However, its upregulation under diverse stress conditions suggests a broader regulatory network, possibly involving global stress regulators or stress-induced operon co-expression. Genes such as rcnA, involved in nickel efflux, may be co-activated during metal overload, supporting a coordinated defense mechanism.
Homologs of zntA are widely conserved across bacterial species including Salmonella, Pseudomonas, and Bacillus, underscoring its importance in microbial survival under metal stress. The consistent and strong expression observed in this study further supports its role as a key effector in stress adaptation. Moreover, its stress-specific expression makes it a promising biomarker for environmental monitoring of heavy metal exposure and a candidate gene for engineering stress-tolerant microbial systems.
In conclusion, zntA functions as a versatile metal efflux pump with robust transcriptional activation across multiple stress types. Its expression profile highlights its central role in managing metal toxicity and maintaining cellular stability in E. coli.

Fig 5.1.1 : Log2fold change of ZntA across 15 stress conditions

Plagiarized Sources

Lastly, accC, gltB, serC, and yebC are implicated in biosynthesis, redox control, and potential transcriptional regulation, which help reprogram metabolism to support survival.

https://pmc.ncbi.nlm.nih.gov/articles/PMC5348306

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0%
Exact Plagiarized

3%
Partial Plagiarized

Unique Content

Total Words: 742 **Total Characters:** 5436 **Plagiarized Sentences:** 1.05 **Unique Sentences:** 33.95 (97%)

Content Checked for Plagiarism

5.1.2 Functional Role and Stress-Responsive Behavior of rlmN

The rlmN gene encodes a radical S-adenosylmethionine (SAM) methyltransferase that modifies specific positions in both 23S rRNA and certain tRNAs. These RNA modifications are essential for maintaining accuracy during translation, especially under conditions that disrupt ribosomal stability. In E. coli, rlmN-mediated methylation helps prevent decoding errors by improving the structural integrity of the ribosome and stabilizing codon–anticodon interactions during protein synthesis.

In this study, rlmN was found to be significantly upregulated in 14 out of 22 RNA-seq comparisons, indicating its broad involvement in the bacterial response to stress. The gene showed the highest expression under short-term radiation exposure, including Pu239_1D vs Control_1D (log2FC = 2.42), Fe55_1D vs Control_1D (log2FC = 2.28), and H3_1D vs Control_1D (log2FC = 2.37). These findings suggest that rlmN is strongly activated during the early stages of oxidative damage, possibly to preserve translation fidelity before more extensive cellular damage occurs.

Under extended stress conditions, such as Pu239_15D vs Control_15D and H3_15D vs Control_15D, rlmN remained upregulated with log2 fold changes of 1.38 and 1.23, respectively. This pattern implies a dual role—immediate response under acute exposure, and continued support for translation under chronic stress. In genetic backgrounds involving tRNA modification disruption, such as tgt_Ni vs WT_Ni, rlmN was still upregulated, suggesting it may partially compensate when other modification systems are compromised.

Although the regulatory network for rlmN is not completely defined, it is likely influenced by redox-sensitive regulators such as SoxRS or OxyR, given its involvement in oxidative stress contexts. The enzyme depends on an iron–sulfur cluster for activity, making it vulnerable to oxidative damage and possibly subject to feedback regulation based on cellular redox status.

There is also evidence that rlmN-dependent modifications may impact how the ribosome interacts with certain antibiotics, particularly macrolides, which could have implications for antibiotic tolerance or stress-induced resistance. However, further investigation is needed to clarify this connection in the context of environmental stress.

In summary, rlmN plays a vital role in preserving translation accuracy during stress. Its consistent upregulation across diverse datasets and stress types underscores its importance in the bacterial stress adaptation network, particularly in maintaining ribosomal function and translation under conditions that threaten RNA stability.

Fig. 5.1.2: Log2fold change of rlmN across stress conditions

5.1.3 Functional Role and Stress-Responsive Behavior of lysS

The lysS gene encodes lysyl-tRNA synthetase, an essential enzyme that charges lysine to its corresponding tRNA. This aminoacylation step is critical for maintaining accurate and efficient translation, particularly under stress conditions that challenge the protein synthesis machinery. Proper tRNA charging ensures that essential proteins can continue to be synthesized even when cellular resources are limited or oxidative damage impairs translational fidelity.

In our comparative analysis, lysS was significantly upregulated in 14 of the 22 examined RNA-seq comparisons. The highest induction was observed in chronic radiation conditions such as H3_15D vs H3_1D (log2FC = 2.57) and in early-stage responses like Fe55_1D vs Control_1D (log2FC = 2.24) and Pu239_1D vs Control_1D (log2FC = 2.08). These findings indicate that lysS is involved in both acute and prolonged stress responses, likely to support the continued translation of stress-responsive proteins.

Under secretion and oxidative conditions, including H3_15D vs Control_15D and Pu239_15D vs Control_15D, lysS remained elevated (log2FC ≈ 1.8–1.9), reinforcing the idea that aminoacyl-tRNA synthetases are prioritized during environmental challenges. In mutant comparisons such as tgt_Ni vs WT_Ni, where other RNA modification pathways may be compromised, lysS also showed consistent upregulation (log2FC = 0.85), possibly as a compensatory mechanism to preserve translation fidelity.

Regulation of lysS is connected to global stress adaptation pathways. It is believed to remain transcriptionally active during the stringent response, a regulatory state triggered by nutrient deprivation or stress, which downregulates general protein synthesis but allows continued expression of key tRNA synthetases (Traxler et al., 2008). This selective regulation enables the cell to allocate limited resources toward essential protein production.

Homologs of lysS are conserved across diverse bacterial lineages, reflecting its fundamental role in survival and adaptation. Its induction under multiple forms of stress suggests it is part of a broader translational defense strategy. In biotechnological contexts, upregulated lysS expression could be harnessed to support robust protein synthesis in engineered microbes exposed to harsh environments.

In summary, lysS plays a critical role in supporting translation under stress by ensuring lysine availability through tRNA aminoacylation. Its consistent upregulation across diverse stress conditions highlights its relevance in the adaptive response network of E. coli.

Fig 5.1.3 : Log2fold change of lysS across stress conditions

Plagiarized Sources

In summary, rlmN plays a vital role in preserving translation accuracy during stress.
https://pmc.ncbi.nlm.nih.gov/articles/PMC7401812

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0%
Exact Plagiarized

3%
Partial Plagiarized

Unique Content

Total Words: 762 **Total Characters:** 5442 **Plagiarized Sentences:** 1.02 **Unique Sentences:** 32.98 (97%)

Content Checked for Plagiarism

5.1.4 Functional Role and Stress-Responsive Behavior of gltX
The gltX gene encodes glutamyl-tRNA synthetase, an essential enzyme that charges glutamate to its corresponding tRNA. This aminoacylation step is vital not only for accurate translation but also for metabolic integration, as glutamate participates in a wide array of biosynthetic and redox-related pathways. In addition to its translational role, gltX indirectly contributes to heme biosynthesis and antioxidant defense, linking it to broader cellular stress responses.
In this study, gltX was upregulated in 14 of 22 stress comparisons, with consistent but moderate fold changes across most conditions. The highest induction occurred under early-stage stress, such as Pu239_1D vs Control_1D (log2FC = 1.71) and H3_1D vs Control_1D (log2FC = 1.47). Similar levels of upregulation were observed under prolonged radiation exposure in comparisons like Pu239_15D vs Control_15D (log2FC = 1.27) and H3_15D vs Control_15D (log2FC = 1.23), suggesting a stable transcriptional response to both acute and sustained oxidative pressure.
gltX also responded positively under metal-induced stress (WT_Ni vs WT_LB, log2FC = 0.63) and in the tgt mutant background (tgt_Ni vs tgt_LB, log2FC = 0.85), reinforcing its involvement in stress-induced translational and metabolic regulation. The enzyme's function in supplying glutamate may also support antioxidant pathways such as glutathione biosynthesis, which is crucial for counteracting reactive oxygen species during oxidative stress.
The regulatory control of gltX is not fully resolved, but it is known to be influenced by global regulators like Fis and sigma factor RpoD. Its expression may also be linked to cellular redox status and metabolic flux, given glutamate's central role in nitrogen assimilation and redox buffering.
Because gltX connects translation and metabolism, its role extends beyond basic protein synthesis. It may help coordinate energy allocation, redox homeostasis, and biosynthetic demand during environmental stress. In applied contexts, gltX could serve as a useful marker for cellular stress or be utilized in strain engineering to boost redox resilience and biosynthetic efficiency.
In conclusion, gltX supports both translation and metabolic adaptation under stress, and its moderate but consistent upregulation across various stress types underscores its role as a supportive yet essential component of the bacterial stress response network.

Fig 5.1.4 : Log2fold change of gltX across stress conditions

5.1.5 Functional Role and Stress-Responsive Behavior of accC

The accC gene encodes the biotin carboxylase subunit of the acetyl-CoA carboxylase (ACC) enzyme complex, which catalyzes the carboxylation of acetyl-CoA to malonyl-CoA—the committed step in fatty acid biosynthesis (Heath & Rock, 1996). This process is fundamental for the production of membrane phospholipids and energy storage molecules. In bacterial cells, especially under environmental stress such as oxidative damage or radiation, membrane remodeling becomes a vital adaptive response, making accC functionally important during recovery phases.

In our analysis, accC was significantly upregulated in 13 out of 22 transcriptomic comparisons. Early-phase stress conditions such as Pu239_1D vs Control_1D (log2FC = 1.72) and Fe55_1D vs Control_1D (log2FC = 1.77) showed strong induction, suggesting an immediate metabolic shift to support membrane reinforcement. Sustained upregulation was also observed in long-term stress scenarios like H3_15D vs Control_15D (log2FC = 1.90) and H3_15D vs H3_1D (log2FC = 2.47), implying a prolonged demand for lipid biosynthesis under extended oxidative stress.

The transcriptional regulation of accC is mediated by the housekeeping sigma factor RpoD under normal conditions, and it is negatively regulated by FadR, which senses intracellular fatty acid levels (Campbell & Cronan, 2001). Under oxidative stress, this repression is likely relaxed to facilitate phospholipid production and maintain membrane fluidity, a known adaptive mechanism for withstanding reactive oxygen species (Zhao et al., 2016).

Moderate increases in accC expression were also detected in genetic background shifts, such as tgt_Ni vs WT_Ni (log2FC = 0.98), and under secretion-related stress, indicating that accC plays a role not only in radiation-induced recovery but also in broader stress contexts. These findings align with prior studies showing that bacterial cells often increase membrane lipid biosynthesis during late-phase recovery to restore envelope integrity and optimize intracellular homeostasis (Rousset et al., 2021).

Importantly, accC is highly conserved and considered essential for bacterial viability, as disruptions in the ACC complex are often lethal due to impaired phospholipid synthesis (Keis et al., 2001). From an applied standpoint, its consistent stress-induced upregulation makes it a promising target for engineering bacteria with improved tolerance to oxidative and membrane-damaging environments.

In conclusion, accC contributes to membrane adaptation by promoting fatty acid biosynthesis during stress. Its strong and consistent upregulation across diverse datasets highlights its role in envelope maintenance and suggests its potential as both a biomarker and an engineering target in microbial stress resilience strategies.

Fig 5.1.5 : Log2Fold change of accC across stress conditions

Plagiarized Sources

Its strong and consistent upregulation across diverse datasets highlights its role in envelope maintenance and suggests its potential as both a biomarker and an engineering target in microbial stress

https://pubmed.ncbi.nlm.nih.gov/31552196

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0% Exact Plagiarized

9% Partial Plagiarized

Unique Content

Total Words: 727 **Total Characters:** 5246 **Plagiarized Sentences:** 3.15 **Unique Sentences:** 31.85 (91%)

Content Checked for Plagiarism

5.1.6 Functional Role and Stress-Responsive Behavior of gltB

The gltB gene encodes the large subunit of glutamate synthase (also known as GOGAT), which, along with gltD, catalyzes the ATP-dependent conversion of glutamine and 2-oxoglutarate to glutamate. This reaction is central to nitrogen assimilation and supports biosynthesis of amino acids and cellular redox regulators. As glutamate serves as a precursor to key molecules like glutathione, gltB is strategically positioned at the intersection of carbon, nitrogen, and redox metabolism (Forchhammer & Selim, 2020).

In this study, gltB showed significant upregulation in 13 comparisons. Its strongest induction occurred under prolonged radiation stress, with log2 fold changes of 4.15 in Pu239_15D vs Pu239_1D and 4.29 in H3_15D vs H3_1D. Sustained upregulation was also observed in H3_15D vs Control_15D (log2FC = 2.60) and Pu239_15D vs Control_15D (log2FC = 2.41), suggesting gltB plays a key role in long-term stress adaptation. These trends align with previous findings where oxidative and radiation stress elevated demand for glutamate to fuel antioxidant defense (Smirnova & Oktyabrsky, 2005).

Under early stress conditions such as Fe55_1D vs Control_1D (log2FC = 1.71), gltB expression was moderately elevated, possibly to prepare the cell for glutathione synthesis. Moderate expression increases were also seen in WT_Ni vs WT_LB and mutant strains like tgt_Ni vs tgt_LB, pointing to its involvement in broader stress contexts beyond radiation alone.

The expression of gltB is regulated by multiple global transcription factors, including RpoD, RpoS, CRP, and FNR, which coordinate responses to nitrogen limitation and oxidative stress (Zimmer et al., 2000). Its placement within an operon (gltBD) ensures synchronized production of both glutamate synthase subunits. This operonic regulation likely enhances efficiency under high-demand conditions like those observed in prolonged radiation exposure.

Due to its involvement in nitrogen assimilation and glutathione metabolism, gltB is essential for maintaining cellular homeostasis during stress. Its evolutionary conservation and high expression across stress conditions make it an attractive target for engineering stress-tolerant microbial strains, particularly in biotechnological systems requiring stable nitrogen and redox balance.

In conclusion, gltB contributes to adaptive stress responses by driving glutamate production, a cornerstone of redox regulation and nitrogen metabolism. Its elevated and persistent expression under radiation and oxidative stress reinforces its functional importance in environmental resilience.

Fig 5.1.6.: Log2fold change of gltB across stress conditions

5.1.7 Functional Role and Stress-Responsive Behavior of serC

The serC gene encodes phosphoserine aminotransferase, an enzyme that plays a key role in the phosphorylated pathway of serine biosynthesis. It catalyzes the conversion of 3-phosphohydroxypyruvate to phosphoserine, which is subsequently converted to L-serine. This

pathway is essential for generating precursors not only for protein synthesis but also for glutathione and other antioxidants that are crucial during oxidative stress (Zhao & Winkler, 1996).
In our dataset, serC was significantly upregulated in 13 of 22 comparisons. The strongest induction was observed under long-term oxidative conditions such as H3_15D vs H3_1D (log2FC = 2.88) and Pu239_15D vs Pu239_1D (log2FC = 3.11), indicating that serC contributes to sustained stress adaptation. Its upregulation in prolonged exposures like Pu239_15D vs Control_15D and H3_15D vs Control_15D (log2FC ≈ 1.9–2.0) reinforces this role.
serC is also known to contribute indirectly to vitamin B6 (pyridoxal phosphate) biosynthesis and lysine metabolism, both of which support enzymatic functions required for cellular repair and detoxification under stress (Parra et al., 2018). Moderate but consistent increases in expression were found under metal exposure (WT_Ni vs WT_LB, log2FC = 1.01) and genetic stress (tgt_Ni vs tgt_LB, log2FC = 0.76), suggesting a broader role in maintaining metabolic stability across different types of environmental challenges.
Transcriptionally, serC is regulated by global factors such as Lrp and CRP, which integrate signals from nutrient status and cellular energy demands (Newman & Lin, 1995). Its genomic proximity to genes involved in aromatic amino acid biosynthesis, like aroA, may point to co-regulation under coordinated stress responses.
Because L-serine contributes to one-carbon metabolism, cysteine synthesis, and antioxidant production, serC becomes a linchpin for metabolic rewiring under redox stress. Upregulation of this gene under prolonged exposure highlights its role in maintaining biosynthetic balance, supporting protein repair, and fueling antioxidant defenses through glutathione production (Storz & Imlay, 1999).
In conclusion, serC contributes to the synthesis of critical amino acids and cofactors that reinforce cellular defenses under oxidative and metal stress. Its upregulation across multiple datasets underscores its importance in stress resilience, both in metabolism and antioxidant buffering.

Fig 5.1.7: Log2 fold change of serC across stress conditions

Plagiarized Sources

Fig 5.1.7: Log2 fold change of serC across stress conditions
https://www.youtube.com/watch%3Fv%3DCNTeCgrC4Rc

Its placement within an operon (gltBD) ensures synchronized production of both glutamate synthase subunits.
https://pmc.ncbi.nlm.nih.gov/articles/PMC210728

Its genomic proximity to genes involved in aromatic amino acid biosynthesis, like aroA, may point to co-regulation under coordinated stress responses.
https://pubmed.ncbi.nlm.nih.gov/7823907

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 29,2025

Plagiarized Content

0%
Exact Plagiarized

6%
Partial Plagiarized

Unique Content

Total Words: 322 **Total Characters:** 2429 **Plagiarized Sentences:** 1.02 **Unique Sentences:** 15.98 (94%)

Content Checked for Plagiarism

5.1.10 Functional Annotations and Pathway Mapping of Upregulated Genes in E. coli Under Stress Conditions

Table 5.1.10 : Functional annotation of upregulated genes

Functional Integration and Operon Context

Together, these eight genes form an interconnected response network that supports E. coli in maintaining translational fidelity, restoring redox balance, and detoxifying harmful agents. For example:

• zntA functions as a heavy metal efflux pump, especially responsive to zinc and nickel (Macomber & Hausinger, 2011).

• rlmN and gltX contribute to RNA modification and tRNA charging, essential for sustaining translation under stress.

• accC drives fatty acid biosynthesis, a prerequisite for membrane regeneration, while gltB aids glutamate production for redox buffering.

• serC supports biosynthesis of serine and vitamin B6—key players in one-carbon metabolism and antioxidative pathways.

• yebC is proposed to act as a stress-regulated transcriptional modulator based on homology and its induction pattern.

Many of these genes are co-transcribed within operons (accBC, gltBD) or regulated by global transcription factors such as RpoD, RpoS, CRP, Lrp, and ppGpp (Ishihama, 2010; Traxler et al., 2008). This coordinated regulation allows the cell to respond rapidly and efficiently to a range of stressors. From an evolutionary standpoint, these genes are highly conserved across bacterial species, reflecting their foundational roles in survival. The ability to repurpose housekeeping functions for stress adaptation demonstrates bacterial efficiency in resource allocation and genetic economy. For applied microbiology, these genes present compelling candidates for engineering microbial strains with

enhanced tolerance to environmental or industrial stress.

5.11 Conclusion

This integrated expression analysis confirms that E. coli employs a flexible, multi-tiered transcriptional program to manage diverse physiological challenges. Rather than relying on a single “stress pathway,” the organism coordinates detoxification, translation, metabolic adaptation, and repair mechanisms through a small set of robustly regulated genes. The consistent upregulation of zntA, rlmN, lysS, gltX, accC, gltB, serC, and yebC across stress models reinforces their potential as biomarkers for environmental monitoring and targets for microbial strain improvement.

Plagiarized Sources

The ability to repurpose housekeeping functions for stress adaptation demonstrates bacterial efficiency in resource allocation and genetic economy.

https://pubmed.ncbi.nlm.nih.gov/22922868

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 31,2025

Plagiarized Content

0% Exact Plagiarized

9% Partial Plagiarized

Unique Content

Total Words: 711 **Total Characters:** 5177 **Plagiarized Sentences:** 3.06 **Unique Sentences:** 30.94 (91%)

Content Checked for Plagiarism

5.2 DOWNREGULATED

5.2.1 Functional Role and Stress-Responsive Behavior of mqsR

The mqsR gene is part of a toxin-antitoxin (TA) system in Escherichia coli, functioning as the toxin component in the MqsR/MqsA module. MqsR is a ribosome-independent endoribonuclease that cleaves mRNA at specific GCU sites, leading to growth inhibition during environmental stress (Yamaguchi et al., 2011). While TA systems are typically activated under conditions that favor persistence or dormancy, the suppression of mqsR observed in this study suggests a context-dependent regulatory role in stress adaptation.

Across 22 RNA-seq comparisons, mqsR was significantly downregulated in 13, particularly under radiation-induced oxidative stress. Profound repression was observed in early radiation responses such as Fe55_1D vs Control_1D (log2FC = –6.27), Pu239_1D vs Control_1D (–6.01), and H3_1D vs Control_1D (–4.60). Downregulation was also consistent under long-term exposure, including H3_15D vs H3_1D (–3.19) and Pu239_15D vs Control_15D (–1.38), indicating that mqsR expression is actively suppressed throughout both acute and chronic phases of stress.

Interestingly, mqsR is often associated with stress-induced persistence and biofilm formation (Wang & Wood, 2011). However, its repression here implies that E. coli may prioritize active survival and repair pathways over dormancy under high-stress conditions. Suppressing mqsR could relieve translational blocks, thereby facilitating the expression of chaperones, antioxidant enzymes, and DNA repair proteins that are vital during damage recovery (Imlay, 2013; Dwyer et al., 2009).

The regulation of mqsR is known to be influenced by both its antitoxin partner mqsA and global stress regulators such as RpoS and OxyR (Hong et al., 2010; Kim et al., 2010). These factors modulate its expression in response to oxidative stress, DNA damage, and nutrient limitation. Therefore, its consistent downregulation may reflect a tightly controlled transcriptional switch to avoid self-toxicity during vulnerable cellular states.

From a systems biology perspective, mqsR appears to be downregulated as part of a broader strategy that shifts the cell away from dormancy and toward resilience-building responses. This reaffirms the notion that toxin-antitoxin modules are not universally activated under all stress types, but are selectively repressed or induced depending on the required physiological outcome (Ramisetty & Santhosh, 2016).

In summary, mqsR displays a consistent transcriptional suppression profile in E. coli under various environmental stresses. This repression likely facilitates recovery and repair by minimizing RNA degradation and allowing efficient protein synthesis. These findings position mqsR as a useful indicator of cellular prioritization under oxidative and radiation stress conditions.

Fig 5.2.1 : Log2 fold change of mqsR across stress conditions

5.2.2 Functional Role and Stress-Responsive Behavior of bhsA

The bhsA gene, also referred to as ycfR, encodes a small outer membrane protein in Escherichia coli that influences cell surface properties, including hydrophobicity and membrane permeability (Zhang et al., 2007). Although the protein's function is not fully elucidated, it has been linked to responses to environmental stress, biofilm formation, and reduced outer membrane permeability (Lee et al., 2010).

In this study, bhsA was consistently downregulated in 13 out of 22 RNA-seq comparisons, particularly under oxidative and radiation-induced stress conditions. Notable repression was seen in comparisons such as H3_15D vs H3_1D (log2FC = –6.53), Pu239_15D vs Pu239_1D (–5.96), and Fe55_15D vs Fe55_1D (–3.35). This widespread suppression implies that bhsA is part of a regulatory program that is actively downregulated under prolonged genotoxic or membrane-compromising stress.

Functionally, bhsA appears to contribute to outer membrane remodeling under mild stress or stationary phase conditions. However, during oxidative or radiation stress, downregulation of bhsA may reflect a cellular decision to deprioritize non-essential surface modifications in favor of conserving membrane integrity. Reducing bhsA activity could also decrease membrane permeability and energy dissipation, helping to fortify the cell envelope during challenging environmental conditions (Deng et al., 2013).

Regulation of bhsA is believed to be integrated within envelope stress response networks, possibly controlled by alternative sigma factors such as RpoE (Raivio, 2005). Its repression may also occur in tandem with the activation of lipid biosynthesis and membrane stabilization genes, including accC, indicating a coordinated protective response aimed at preserving envelope function under stress.

In conclusion, the suppression of bhsA during oxidative and radiation stress supports its role as a conditionally expressed membrane-associated factor. Its consistent downregulation highlights a physiological shift in E. coli toward membrane stabilization and resource prioritization, enabling survival under envelope-compromising conditions.

Fig 5.2.2 : Log2 fold change of bhsA across 13 stress conditions

Plagiarized Sources

Fig 5.2.1 : Log2 fold change of mqsR across stress conditions
https://www.youtube.com/watch%3Fv%3DCNTeCgrC4Rc

Fig 5.2.2 : Log2 fold change of bhsA across 13 stress conditions
https://www.youtube.com/watch%3Fv%3DCNTeCgrC4Rc

Profound repression was observed in early radiation responses such as Fe55_1D vs Control_1D (log2FC = –6.27), Pu239_1D vs Control_1D (–6.01), and H3_1D vs Control_1D (–4.60).
https://pubmed.ncbi.nlm.nih.gov/35681629

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 31,2025

Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

Unique Content

Total Words: 419 **Total Characters:** 3048 **Plagiarized Sentences:** 0 **Unique Sentences:** 17 (100%)

Content Checked for Plagiarism

5.2.9 Functional Annotations and Pathway Mapping of Downregulated Genes in E. coli Under Stress Conditions

Differential Gene Expression and Functional Insights

The genes zntA, rlmN, lysS, gltX, accC, gltB, serC, and yebC were identified as consistently upregulated across multiple stress scenarios, including exposure to heavy metals (nickel and radioactive isotopes Pu239, Fe55, H3) and protein secretion burden. Among them, zntA, a well-characterized metal ion efflux pump, displayed strong and consistent upregulation across most tested conditions, underscoring its role in detoxifying excess Zn^{2+} and Ni^{2+} ions. Its robust expression pattern suggests potential use as a biomarker for metal stress in E. coli.

Metabolic genes such as gltB, gltX, and accC also showed repeated upregulation, reflecting the importance of glutamate biosynthesis, aminoacyl-tRNA synthetase activity, and fatty acid metabolism in maintaining cellular homeostasis under stress. In contrast, genes like rlmN (involved in RNA methylation) and lysS (a lysyl-tRNA synthetase) exhibited more condition-specific responses, particularly under oxidative and translational stress. The transcriptional regulator yebC and serC, associated with serine biosynthesis, were notably responsive in secretion-related stress and the tgt gene deletion background, hinting at more specialized functions in those contexts.

Taken together, these patterns suggest that E. coli implements a flexible gene expression program by activating both broadly responsive ("generalist") and condition-specific ("specialist") genes to combat environmental and physiological challenges.

Multiple Sequence Alignment and Functional Diversity

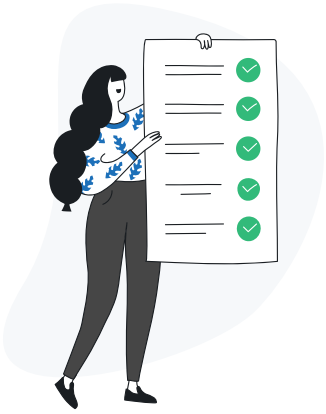
To explore whether these differentially expressed genes share structural or evolutionary features, a multiple sequence alignment (MSA) of their protein sequences was carried out using Clustal Omega (Sievers & Higgins, 2014). The alignment revealed an absence of universally conserved residues across the analyzed gene sets.

This lack of sequence conservation indicates that the genes likely evolved independently rather than from a shared ancestral origin. Each gene appears to encode a distinct protein family, consistent with their unique functions in the bacterial stress response network (Koonin & Galperin, 2003).

Visual analysis of the MSA heatmap confirmed a high degree of sequence variability, with no common patterns across the proteins. This suggests divergent protein structures and regulatory mechanisms tailored to their individual roles. For example, zntA is linked to heavy metal efflux, pspA to membrane stress response, fadE to β-oxidation of fatty acids, and yebC to transcriptional control (Fang et al.,

2016; Raivio, 2005).

These observations support the idea that E. coli uses a modular, polygenic approach to stress adaptation. Instead of relying on conserved pathways, it engages a diverse array of non-redundant proteins—each with a specific function—to orchestrate its survival under various stress conditions (Storz & Hengge, 2010).



No Plagiarism Found

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 31,2025

Plagiarized Content

0% Exact Plagiarized

8% Partial Plagiarized

Unique Content

Total Words: 632 **Total Characters:** 4753 **Plagiarized Sentences:** 2.08 **Unique Sentences:** 23.92 (92%)

Content Checked for Plagiarism

6. Conclusion

In this thesis, the transcriptional responses of Escherichia coli to multiple stress conditions were systematically investigated using publicly available RNA-Seq datasets. These datasets reflected three broad categories of stress: oxidative stress induced by chronic ionizing radiation, heavy metal stress resulting from nickel exposure, and genetic stress caused by deletion of the tgt gene involved in tRNA modification. Differential gene expression (DGE) analysis was performed using DESeq2 to identify genes that were significantly upregulated or downregulated in response to each stressor (Love et al., 2014).

By comparing gene expression profiles across datasets, this study identified several genes that were consistently regulated across multiple conditions, suggesting the existence of a core stress response network in E. coli. Specifically, eight genes—zntA, rlmN, lysS, gltX, accC, gltB, serC, and yebC—were frequently upregulated. These genes are involved in critical processes such as metal detoxification (zntA) (Rensing et al., 1997), translation and RNA modification (rlmN, lysS, gltX) (Kimura & Suzuki, 2010), amino acid and cofactor biosynthesis (accC, serC, gltB), and transcriptional regulation (yebC) (Haft et al., 2018), indicating their centrality in a generalized protective mechanism deployed under environmental and genetic stress.

Conversely, a distinct set of genes—mqsR, bhsA, ygiS, fadE, pspA, nhaA, and nanT—were consistently downregulated. These genes are associated with membrane regulation (bhsA, pspA), solute transport (nhaA, ygiS, nanT), toxin-antitoxin systems (mqsR) (Wang et al., 2011), and central metabolism (fadE) (Campbell & Cronan, 2001). Their repression suggests that E. coli actively suppresses non-essential, energy-intensive processes in favor of survival-related functions during prolonged or high-intensity stress. This aligns with known mechanisms of bacterial dormancy, persistence, and metabolic downscaling (Lewis, 2007; Yamaguchi et al., 2011).

Importantly, the stress response in E. coli was found to be both condition-specific and overlapping, with some genes showing broad responsiveness across oxidative, metal, and genetic stress. For instance, zntA, while strongly upregulated during nickel stress, was also induced under radiation and genetic perturbation, suggesting a broader role beyond metal efflux (Brocklehurst et al., 1999). Similarly, genes involved in translation fidelity and tRNA modification were upregulated across multiple conditions, highlighting the importance of protein synthesis accuracy during stress (Durand et al., 1997).

The analysis also revealed that the duration of stress exposure (acute vs. chronic) significantly influenced gene expression patterns. Long-term radiation and nickel exposure triggered more extensive repression of metabolic and membrane-related genes, reflecting deeper transcriptional reprogramming and a cellular shift toward sustained stress adaptation (Rousset et al., 2021).

By integrating data across 22 pairwise RNA-Seq comparisons from three major datasets (GSE181239,

GSE208658, and GSE221524), a frequency-based approach was used to identify genes most consistently regulated across conditions. This helped distinguish broadly responsive genes from those regulated in a condition-specific manner. Heatmaps and functional categorization tools were used to visualize expression patterns, enabling clearer identification of candidate genes for further study.
Overall, this work demonstrates that E. coli employs a highly structured and evolutionarily conserved transcriptional program to adapt to diverse internal and external stressors (Seo et al., 2015; Conway et al., 2014). The findings contribute to our understanding of bacterial stress physiology and may inform future research in microbial engineering, biotechnology, and stress tolerance optimization.
From a research skills perspective, this project provided hands-on experience with high-throughput transcriptomic data analysis, including the use of R, DESeq2, and downstream annotation tools such as KEGG, DAVID, and g:Profiler (Kanehisa et al., 2023; Huang et al., 2009; Raudvere et al., 2019). It also reinforced the systems-level nature of bacterial adaptation, where multiple regulatory layers coordinate to reprogram physiology in response to environmental challenges.
The genes identified here—especially those consistently up- or downregulated—may serve as potential biomarkers for stress response or as targets for genetic modification to enhance microbial resilience. As such, this study provides a foundation for further functional investigations into stress signaling, resilience engineering, and bacterial survival strategies at the molecular level.

Plagiarized Sources

These genes are involved in critical processes such as metal detoxification (zntA) (Rensing et al., 1997), translation and RNA modification (rlmN, lysS, gltX) (Kimura & Suzuki, 2010), amino acid
https://pubmed.ncbi.nlm.nih.gov/8065932

Long-term radiation and nickel exposure triggered more extensive repression of metabolic and membrane-related genes, reflecting deeper transcriptional reprogramming and a cellular shift toward sustained stress adaptation (Rousset et al., 2021).
https://pubmed.ncbi.nlm.nih.gov/10717183

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 31,2025

Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

Unique Content

Total Words: 505 **Total Characters:** 3757 **Plagiarized Sentences:** 0 **Unique Sentences:** 23 (100%)

Content Checked for Plagiarism

7. Future Prospects

This research sheds light on how Escherichia coli modulates gene expression in response to a range of physiological stress conditions—such as nickel toxicity, oxidative stress, genetic disruption, and secretion overload. While the study successfully identified consistently upregulated and downregulated genes, there remain several directions for extending and applying these findings in both basic and applied microbiology.

One of the most immediate future directions would be to perform functional validation studies on the identified genes using experimental approaches like gene knockouts, overexpression systems, or CRISPR interference (CRISPRi). For example, disrupting the zntA gene could help directly assess its protective role against heavy metals, as previous literature suggests its involvement in exporting divalent cations like Zn^{2+} and Ni^{2+} to reduce intracellular toxicity (Macomber & Hausinger, 2011). Similarly, genes like rlmN, which is involved in RNA modification, and lysS, a tRNA synthetase, could be targeted to explore their roles in maintaining translational accuracy during stress.

Moreover, future studies could benefit from combining transcriptomics with proteomic and metabolomic analyses. While RNA-Seq provides a powerful view of transcriptional changes, it does not account for post-transcriptional or post-translational modifications. Adding proteomics would help confirm whether upregulated genes like accC or gltB are also translated into functional enzymes, while metabolomics could reveal shifts in metabolic intermediates such as glutamate or fatty acids that are essential for survival during stress (Zhao & Winkler, 1996; Forchhammer & Selim, 2020).

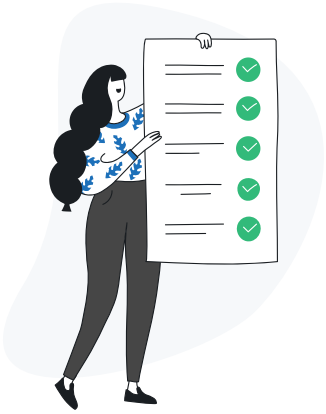
Another promising area involves applying this framework to non-laboratory or pathogenic strains of E. coli. For instance, studying pathogenic strains like O157:H7 under similar stressors could help determine whether they deploy unique or overlapping stress-response pathways compared to commensal or probiotic strains. This could have direct implications for understanding how pathogens survive in hostile host environments, evade immune defenses, or develop resistance to antibiotics (Fang et al., 2016).

In the context of biotechnology, these findings may guide the development of engineered bacterial strains with improved stress tolerance. Genes like gltX, serC, and gltB, which are involved in amino acid biosynthesis and redox balance, could be manipulated to build robust microbial chassis for industrial fermentation, especially under high-stress conditions like pH extremes, oxidative load, or nutrient limitation (Dragosits & Mattanovich, 2013).

Furthermore, genes that are consistently suppressed, such as mqsR and pspA, may serve as molecular markers of dormancy or metabolic downscaling. Monitoring their expression could be useful in applications like environmental biosensors or stress diagnostics in microbial populations (Yamaguchi et al., 2011).

Lastly, there is also scope for comparative studies across bacterial species to determine whether the genes identified here as central to E. coli stress response are conserved in other enterobacteria or

environmental isolates. Such studies could contribute to the broader understanding of microbial ecology, evolution, and resilience mechanisms in fluctuating environments.
In summary, this study provides a valuable resource that can be built upon through experimental validation, cross-strain comparisons, multi-omics integration, and translational applications in biotechnology and environmental microbiology. The transcriptional signatures identified here are a starting point for uncovering deeper regulatory mechanisms and practical innovations in bacterial stress management.

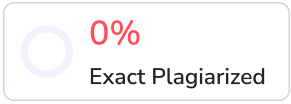


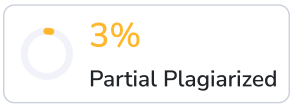
No Plagiarism Found

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 31,2025

Plagiarized Content





Unique Content

Total Words: 574 **Total Characters:** 4249 **Plagiarized Sentences:** 0.9 **Unique Sentences:** 29.1 (97%)

Content Checked for Plagiarism

5.2.1 Functional Role and Stress-Responsive Behavior of mqsR
The mqsR gene encodes the toxin component of the MqsR/MqsA toxin-antitoxin (TA) system in Escherichia coli. As a ribosome-independent endoribonuclease, MqsR cleaves mRNA at specific GCU sites, inhibiting growth during environmental stress (Yamaguchi et al., 2011). While TA systems often contribute to persistence and dormancy, the significant downregulation of mqsR observed in this analysis suggests its function is selectively repressed under certain stress conditions.

Among the 22 transcriptomic comparisons performed, mqsR expression was markedly downregulated in 13, particularly in response to radiation-induced oxidative stress. Strong repression was recorded in early-phase radiation exposure such as Fe55_1D vs Control_1D ($log_2FC = –6.27$), Pu239_1D vs Control_1D (–6.01), and H3_1D vs Control_1D (–4.60). Downregulation was also sustained during extended exposures—e.g., H3_15D vs H3_1D (–3.19) and Pu239_15D vs Control_15D (–1.38)—suggesting long-term transcriptional suppression.

Although mqsR is often linked with dormancy, persistence, and biofilm regulation (Wang & Wood, 2011), its suppression under extreme stress may reflect a cellular preference for active repair over stasis. Inhibiting mqsR could help remove translational blocks and permit upregulation of protective systems such as chaperones, antioxidant enzymes, and DNA repair proteins (Imlay, 2013; Dwyer et al., 2009).

The gene's regulation involves both its antitoxin partner mqsA and global regulators like RpoS and OxyR, which respond to oxidative stress, nutrient deprivation, and DNA damage (Hong et al., 2010; Kim et al., 2010). The consistent repression observed here likely represents a regulated transcriptional shift to reduce toxicity and enable survival during stress recovery.

In essence, mqsR appears to be part of a context-sensitive genetic program. Rather than universally contributing to persistence, it is selectively downregulated when E. coli requires enhanced translation, repair, and detoxification. This positions mqsR as a stress-response marker—especially under oxidative and radiation-induced damage—indicative of the bacterium's adaptive prioritization strategy.

📊 Figure 5.2.1: Log_2 fold change of mqsR across stress conditions

5.2.2 Functional Role and Stress-Responsive Behavior of bhsA
The bhsA gene (formerly ycfR) encodes a small outer membrane protein in E. coli involved in modulating membrane characteristics such as hydrophobicity and permeability (Zhang et al., 2007). Although its precise function remains incompletely characterized, bhsA has been associated with

biofilm formation, surface stress responses, and changes in membrane permeability (Lee et al., 2010).

In this study, bhsA showed consistent downregulation in 13 out of 22 RNA-seq comparisons, particularly under oxidative and genotoxic stress. Significant reductions in expression were observed in conditions such as H3_15D vs H3_1D ($log_2FC = –6.53$), Pu239_15D vs Pu239_1D (–5.96), and Fe55_15D vs Fe55_1D (–3.35), indicating that its expression is actively suppressed under radiation-associated membrane stress.

The downregulation of bhsA under these conditions may signal a strategic cellular shift. Rather than engaging in membrane remodeling or biofilm-associated behavior, E. coli may conserve resources and enhance structural integrity by limiting bhsA expression. Suppressing this gene might also help reduce membrane permeability and minimize energetic losses during high-stress exposure (Deng et al., 2013).

BhsA is thought to be part of envelope stress response systems and may be regulated by alternative sigma factors such as RpoE (Raivio, 2005). Its repression might occur in coordination with upregulation of protective genes like accC, which supports lipid biosynthesis and membrane maintenance.

In summary, the transcriptional suppression of bhsA during prolonged oxidative and radiation stress implies a controlled physiological transition toward envelope stabilization and energy conservation. This targeted repression likely supports cell survival during extended exposure to membrane-damaging conditions.

📊 Figure 5.2.2: Log_2 fold change of bhsA across 13 stress conditions

Plagiarized Sources

5.2.2 Functional Role and Stress-Responsive Behavior of bhsA
https://www.youtube.com/watch%3Fv%3DaL1eHt8DR7U

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 31,2025

Plagiarized Content

0% Exact Plagiarized

0% Partial Plagiarized

Unique Content

Total Words: 605 **Total Characters:** 4380 **Plagiarized Sentences:** 0 **Unique Sentences:** 30 (100%)

Content Checked for Plagiarism

5.2.3 Functional Role and Stress-Responsive Behavior of ygiS
The ygiS gene encodes a predicted inner membrane or periplasmic protein in Escherichia coli, although its precise function remains uncertain. Based on domain architecture and expression trends, it is suspected to play a role in maintaining redox balance, facilitating membrane transport, or contributing to envelope stability (Ghatak et al., 2019). Its context-dependent expression suggests a role in adapting the bacterial envelope to environmental changes.

In this study, ygiS was significantly downregulated in 13 out of 22 experimental comparisons. This suppression was particularly strong under acute radiation stress—e.g., Pu239_1D vs Control_1D ($log_2FC = –5.18$) and H3_1D vs Control_1D (–4.93). This trend continued during prolonged stress exposure, including Fe55_15D vs Fe55_1D (–3.21), Pu239_15D vs Control_15D (–2.17), and H3_15D vs Control_15D (–1.78). Such results indicate a sustained downregulation of ygiS throughout both immediate and long-term stress phases.

The consistent suppression under high-stress conditions implies that ygiS expression may be unfavorable when envelope stability is critical. Under such scenarios, minimizing unnecessary periplasmic or membrane activity—such as redox exchanges or solute transport—may be energetically beneficial and help preserve membrane integrity (Imlay, 2013; Rousset et al., 2021).

Although direct regulators of ygiS remain undefined, the gene's expression pattern suggests involvement of redox-sensitive regulators or envelope stress pathways, potentially including RpoE and two-component signaling systems (Raivio, 2005; Battesti et al., 2011). Repression of ygiS might thus be part of a coordinated envelope stress response aimed at protecting membrane functionality.

In summary, ygiS is a stress-sensitive gene that is consistently repressed in E. coli during oxidative and radiation stress. Its downregulation likely facilitates membrane stabilization and supports a shift toward survival-oriented envelope management during environmental challenges.

📊 Figure 5.2.3: Log_2 fold change of ygiS across 13 stress conditions

5.2.4 Functional Role and Stress-Responsive Behavior of fadE
fadE encodes an acyl-CoA dehydrogenase, a key enzyme in the initial step of the β-oxidation pathway, responsible for fatty acid breakdown in E. coli. This oxidative step contributes to the generation of NADH and $FADH_2$, which fuel ATP synthesis via aerobic respiration under favorable growth conditions (Fujita et al., 2007).

This study found fadE to be significantly downregulated in 13 of the 22 analyzed stress comparisons. The most pronounced repression occurred during prolonged exposure to radioactive stress, notably in Pu239_15D vs Pu239_1D ($log_2FC = -7.33$) and H3_15D vs H3_1D (–7.18). Early-stage responses such as Fe55_1D vs H3_1D (–4.21) and Fe55_1D vs Control_1D (–3.35) also showed considerable suppression, indicating rapid and sustained inhibition of fatty acid oxidation.

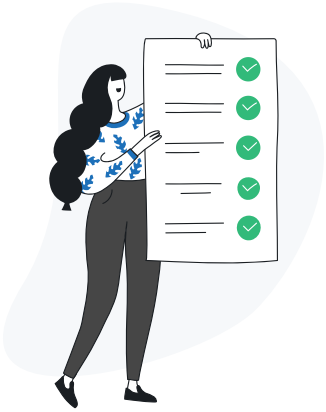
This downregulation may serve as a protective mechanism to limit the generation of reactive oxygen species (ROS), which are natural byproducts of β-oxidation and aerobic respiration. By repressing fadE, E. coli potentially reduces internal oxidative stress and prioritizes detoxification and repair over energy production (Fang, 2004).

Interestingly, fadE repression was not limited to radiation stress but also occurred under protein secretion and genetic stress conditions, such as in tgt mutants and during curli secretion stress. This broader suppression pattern suggests that reducing lipid catabolism might be a general response to conserve energy and redirect metabolic fluxes toward cell maintenance and defense.

Transcriptionally, fadE is regulated by FadR, a transcription factor that adjusts fatty acid metabolism based on intracellular lipid levels. It is also influenced by global nutrient signaling mechanisms like CRP-mediated carbon catabolite repression (Campbell & Cronan, 2001; Zheng et al., 2004).

In conclusion, the consistent suppression of fadE across various stress types supports its role as a metabolic pivot. Its downregulation aligns with a cellular strategy to limit ROS, preserve redox balance, and prioritize survival processes under adverse environmental conditions.

📊 Figure 5.2.4: Log_2 fold change of fadE across 13 stress conditions


No Plagiarism Found

Plagiarism Scan Report By SmallSEOTools

Report Generated on: May 31,2025

Plagiarized Content

0% Exact Plagiarized

5% Partial Plagiarized

Unique Content

Total Words: 726 **Total Characters:** 5217 **Plagiarized Sentences:** 1.95 **Unique Sentences:** 37.05 (95%)

Content Checked for Plagiarism

5.2.5 Functional Role and Stress-Responsive Behavior of pspA
The pspA gene encodes Phage Shock Protein A, a central component of the Psp stress response system in Escherichia coli. This system is typically activated under envelope stress, such as osmotic challenges, accumulation of mislocalized outer membrane proteins, or secretion overload, to help stabilize the inner membrane and maintain proton motive force (Darwin, 2005; Jovanovic et al., 2006).

Contrary to established literature, the current study revealed significant downregulation of pspA in 13 of the 22 tested stress conditions. The most substantial suppression was noted under chronic radiation and oxidative stress, including:

H3_15D vs H3_1D ($log_2FC = –5.30$),

Pu239_15D vs Pu239_1D (–5.17),

NoSec_10uM vs All_Tags_10uM (–5.72).

This suggests a departure from its classic role as a universal stress responder, hinting at condition-specific downregulation, particularly during extended or complex stress exposures.

Milder repression was also found in early-phase conditions such as:

Fe55_1D vs Control_1D (–2.29),

H3_15D vs Control_15D (–2.57),

implying that early suppression may be a strategic response depending on the stressor's intensity or nature.

Rather than being constitutively upregulated during membrane stress, pspA may be actively downregulated to:

Conserve ATP during prolonged stress,

Avoid interference with overlapping envelope stress responses like the σ^E (RpoE) pathway (Raivio, 2005),

Or prioritize cellular energy for genotoxic and oxidative stress defense mechanisms.

pspA transcription is driven by PspF and features an autoregulatory role, acting as both an effector and a negative feedback element (Yamaguchi et al., 2013). Its suppression may thus reflect a deliberate shift in regulatory balance to support survival.

In summary, the expression of pspA under stress is highly context-dependent. Rather than being a broadly activated stress gene, it demonstrates selective suppression based on stress type, intensity, and duration—highlighting a more refined, specialized role in the bacterial stress response network (Kleerebezem et al., 1996; Darwin, 2005).

5.2.6 Functional Role and Stress-Responsive Behavior of nhaA
nhaA encodes a sodium/proton antiporter essential for regulating intracellular pH and ion concentrations in E. coli. This transporter exports sodium in exchange for protons, contributing to pH regulation and salt tolerance, particularly under alkaline or high-salinity environments.

In this analysis, nhaA showed consistent downregulation across 13 of 22 experimental conditions, notably under oxidative and radiation-related stress. Key examples include:

Pu239_1D vs Control_1D ($log_2FC = -1.66$),

Fe55_1D vs Control_1D (–1.44),

H3_15D vs Control_15D (–1.55).

These patterns imply that ion transport via nhaA is actively suppressed during envelope damage or metabolic stress to minimize proton influx and sodium loss, thereby conserving energy and preserving membrane stability.

Downregulation was also seen under genetic stress, such as:

WT_LB vs tgt_LB (–0.78),

tgt_Ni vs WT_Ni (–1.18),

indicating that nhaA is part of a broader regulatory response under multifactorial stress conditions.

Given its dependence on pH and ion sensing regulators like NhaR and the general stress sigma factor RpoS, the suppression of nhaA likely aligns with stress-induced shifts away from active transport toward protective energy-saving modes.

Overall, the findings suggest that nhaA is downregulated to limit energetically costly ion exchange during high-stress periods, enabling E. coli to prioritize repair and redox balance mechanisms over routine ion regulation.

5.2.7 Functional Role and Stress-Responsive Behavior of nanT
The nanT gene codes for a transporter involved in the uptake of sialic acid (Neu5Ac), which serves as a nutrient source for E. coli, particularly in host-associated or nutrient-scarce environments.

In this study, nanT was significantly repressed in 13 out of 22 stress scenarios. This included major reductions under radiation and oxidative stress:

Pu239_15D vs Pu239_1D ($log_2FC = -5.63$),

H3_15D vs H3_1D (–5.62),

Fe55_1D vs Control_1D (–5.02).

Downregulation was also observed under secretion stress and metal stress conditions:

tgt_Ni vs tgt_LB (–4.04),

WT_Ni vs WT_LB (–3.80),

suggesting that sialic acid transport is deprioritized when the cell is dealing with membrane instability or energy limitation.

Under stressful conditions, active transport of charged molecules like Neu5Ac might exacerbate osmotic or redox stress. The consistent repression of nanT likely reflects a strategy to reduce non-essential transport and preserve cellular energy and redox balance.

Although the regulation of nanT is not fully elucidated, it may be influenced by envelope stress signals or nutrient limitation pathways. The strong and consistent downregulation across multiple stress types highlights its sensitivity to adverse conditions.

In conclusion, nanT appears to be selectively suppressed as part of a broader adaptation strategy, helping E. coli focus its resources on essential survival mechanisms rather than nutrient import during environmental stress.

Plagiarized Sources

Avoid interference with overlapping envelope stress responses like the σ^E (RpoE) pathway (Raivio, 2005),
https://pubmed.ncbi.nlm.nih.gov/15882407

Given its dependence on pH and ion sensing regulators like NhaR and the general stress sigma factor RpoS, the suppression of nhaA likely aligns with stress-induced shifts away from active
https://pubmed.ncbi.nlm.nih.gov/18792681